Three-Dimensional Registration and Tracking of Vascular Structures Using Calibrated Biplane Fluoroscopy

by

Michael Hayden Rosenthal

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Science in the Department of Computer Science.

Chapel Hill
2005

Approved by

Advisor: Professor Henry Fuchs

Reader: Professor Fred Brooks

Reader: Professor Elizabeth Bullitt

Reader: Professor Guido Gerig

Reader: Professor Matthew Mauro

Reader: Professor Greg Welch

ABSTRACT

Michael Hayden Rosenthal: Three-Dimensional Registration and Tracking of Vascular Structures Using Calibrated Biplane Fluoroscopy
(Under the direction of Professor Henry Fuchs)

Medicine is rapidly adopting new minimally invasive techniques, especially in the realm of vascular procedures. These new techniques require spatially complex operations while providing little or no hand-eye coordination. Image-based tracking, registration, and visualization may make it easier to understand and perform these procedures, possibly shortening procedures, improving outcomes, and allowing new, more complex techniques to be pursued. This dissertation will explore the calibration, tracking, and registration issues that need to be understood to develop a viable intraoperative guidance system.

To develop an understanding of the likely benefits of a complete system for intraoperative tracking and visualization, this dissertation will consider the methods, accuracy and effectiveness of several major components in such a system. The first part of this work presents a novel phantom for simultaneous calibration of two fluoroscopes, a set of methods for accurate calibration of such systems based on bundle adjustment, and the integration of high-order distortion correction and calibration into a single optimization method. Existing methods generally address single-view calibration and distortion correction as separate steps, with the best fluoroscopic methods producing mean errors of at least 1-2mm. The methods presented herein yield a mean reconstruction error of 0.44 mm from a single calibration process. The second part describes a real-time anatomical feature tracker based on maximum-likelihood feature selection. This method is then used to determine the motion of the liver in fluoroscopic image sets from real patients. The final innovation of this dissertation is a method for registering a 3D vascular model to two calibrated angiographic views simultaneously.

Overall, this work advances our understanding of the ways in which real-time fluoroscopy

can be used to precisely track and model interventional procedures interactively. These results enable sub-millimeter reconstruction of anatomical features, instruments, or other features of interest in a practical and robust manner.

DEDICATION

This work is dedicated to my mother, Georgia Bernstein, who has tirelessly and selflessly supported me throughout my life. She has made innumerable sacrifices to ensure my education and development; I have become the person I am today because of her. If this work leaves a mark on our knowledge as a society, it is a mark drawn by her love, support, and vision throughout my life.

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to his committee for their expert guidance, motivation, and patience. The author would also like to acknowledge Susan Weeks, M.D., Andrei State, M.S., Marc Pollefeys, Ph.D., and Keith Muller, Ph.D. for their support of and contributions to this research. This work was funded in part by NIH-HLB grant 1-R01-HL69808-01. Marc Pollefeys' contribution was partially supported by his NSF Career award IIS-0237533.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D: Two dimensional

3D: Three dimensional

AP: Anteroposterior (front to back), generally referencing the path of travel of the x-rays from source to imaging plane

CT: Computed tomography, a method of 3D imaging based on x-ray absorption

MR: Magnetic resonance, a method of 3D imaging based on electron spin variations

PA: Posterior-to-anterior (back to front), used as AP above. This is the most common orientation for chest imaging.

TIPS: Transjugular Intrahepatic Portosystemic Shunt

GLOSSARY

ANGIOGRAM: An angiogram is an image that emphasizes blood vessels. These images are commonly produced using fluoroscopy, an x-ray imaging method, although other methods can be used. An agent that increases the contrast of the blood vessels relative to surrounding structures is often injected into the blood stream to yield better images.

FLUOROSCOPY: Fluoroscopy is the x-ray equivalent of a television camera. This type of imaging device can be used during an operation to interactively visualize bones, instruments, and certain contrast agents. A "biplane" fluoroscopy system has two fluoroscopy cameras at different positions (stereoscopic vision) and can provide additional detail about the position of objects within a patient.

TIPS: Transjugular Intrahepatic Portosystemic Shunt — a medical procedure to relieve some of the symptoms of advanced liver disease. The procedure involves the creation of a tunnel, or shunt, from the abnormally pressurized portal vein in the bottom of the liver to the draining hepatic vein in the top of the liver. This relieves the excess pressure in the portal vein. See page 4 for additional detail.

CALIBRATION: The process of determining or adjusting parameters of a system or device. In this work, calibration refers to the determination of the geometric parameters of a fluoroscopic imaging system.

2D-3D REGISTRATION: The process of determining the relationship between 2D imaging systems (e.g., fluoroscopes) and some 3D object being imaged.

Part I

Introduction and Background

Chapter 1

INTRODUCTION

Medicine is rapidly adopting new minimally invasive techniques, especially in the realm of vascular procedures. These techniques offer faster recovery times and better quality of life for patients, but they often involve complex procedures that demand precision while offering little or no hand-eye coordination. Image-based tracking, registration, and visualization may make it easier to understand and perform these procedures, possibly shortening procedures, improving outcomes, and allowing new, more complex techniques to be pursued. This dissertation will explore the calibration and real-time image processing issues that need to be understood to develop a safe, reliable, and useful system for intraoperative guidance of these tasks.

The Transjugular Intrahepatic Portosystemic Shunt (TIPS) procedure will be used as a model for exploring these issues. This procedure was chosen due to the obvious need for improved guidance in this task and the possible extension of this technology to many other related procedures. A research group consisting of Elizabeth Bullitt MD, Sue Weeks MD, Stephen Aylward PhD, Henry Fuchs PhD, and myself have been working to provide three-dimensional (3D) guidance and visualization for this procedure. This dissertation will encompass only a part of the group's total research plan.

To develop an understanding of the likely benefits of a complete system for intraoperative tracking and visualization, this dissertation will consider the methods, accuracy and effectiveness of several major components in such a system. Part I introduces the motivation for this research, discusses the state of the art prior to this work, and provides perspective on the overall TIPS guidance system that will assist in understanding key parts of this work.

Part II explores the calibration of biplane fluoroscopy systems. Three-dimensional reconstruction requires knowledge of the geometry of the imaging system, so all of the tracking,

registration, and visualization to be done in this project demands accurate calibration of the biplane fluoroscopy system. Calibration of such systems is complicated by the narrow fields of view, long relative focal lengths, and complex distortion patterns typical of fluoroscopy systems. This work includes the development of a novel phantom for simultaneous calibration of two fluoroscopes, a set of methods for accurate calibration of such systems based on bundle adjustment, and the integration of high-order distortion correction and calibration into a single optimization method. This part closes with an analysis of the performance of the integrated calibration method.

Part III addresses the problem of tracking a low-contrast anatomical feature (the liver) in a complex operative environment. A novel approach to organ tracking – inflating a common intravascular balloon catheter using a radio-opaque contrast agent in the organ of interest – forms the foundation of these methods. This elliptical balloon can be tracked in three dimensions of translation and two angles of rotation, so most of the motion of the liver during a TIPS procedure can be measured. The fiducial is automatically identified and followed using a maximum-likelihood approach that selects the best balloon from among all possible image features. The final chapter in this part describes the motion of the liver in real fluoroscopic patient data using this tracking approach.

The final novel part of this dissertation, Part IV, describes a method for registering a 3D vascular model to two calibrated angiographic views simultaneously (2D-3D registration). This approach leverages the accurate calibration provided in Part II to allow simultaneous registration of the vessel model to both fluoroscopic views. A novel distance-weighting scheme is introduced to normalize errors between views and to correct for perspective scaling effects.

The appendices of this dissertation include descriptions of supporting work by this and other authors. While this material is not core to the dissertation, it is included for the sake of completeness and ease of reference.

Overall, this work advances our understanding of the ways in which real-time fluoroscopy can be used to precisely track and model interventional procedures interactively. The methods and results presented enable sub-millimeter reconstruction of anatomical features, instruments, or other features of interest in a practical and robust manner.

Chapter 2

BACKGROUND

This chapter provides an overview of the medical and technical components of the TIPS procedure. It will close with a discussion of other guidance systems and research of relevance to this work.

2.1 TIPS Procedure

The *Transjugular Intrahepatic Portosystemic Shunt* (TIPS) procedure is used to relieve the life-threatening symptoms of patients with end-stage liver disease. It is also used as a bridge to transplantation for those waiting to receive a new liver. The number of patients with liver failure continues to increase in the United States, with over 12,000 people awaiting liver transplantation in recent years [39]. Many of these patients will undergo TIPS for life-threatening symptoms like esophageal varices (bleeding vessels), excessive, uncontrollable fluid retention within the abdomen (ascites), and associated renal failure (hepatorenal syndrome) or respiratory compromise (hepatopulmonary syndrome).

These symptoms arise when the liver is no longer able to serve its usual role as the body's filter. The liver usually receives nutrient-rich blood from the intestines via the portal vein. It also receives oxygenated blood via the hepatic arteries. The liver effectively filters this blood, processing toxins and excreting undesirable compounds via the biliary system. The processed blood then returns to the systemic circulation via the hepatic veins (see Figure 2.1).

In patients with severe liver disease, the liver acts like a blocked filter and resists the flow of blood. This yields increased pressure in the portal vein that can lead to uncontrollable fluid leakage into the abdomen and severe internal bleeding.

The TIPS procedure creates a direct conduit between the portal vein (going from the

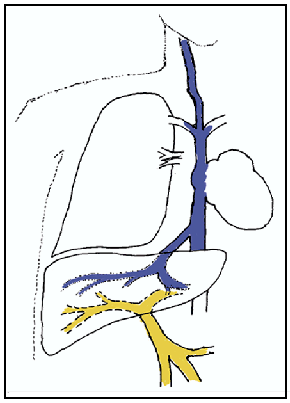


Figure 2.1: A drawing of the liver in relation to vena cava and hepatic veins (blue) and portal veins (yellow)[2].

intestines to the liver) and the hepatic vein (the vein that drains the liver), as shown in Figure 2.2. This conduit allows decompression of elevated portal vein pressures and subsequent relief of symptoms. A radiologist performs the procedure under ultrasound and fluoroscopic guidance. A catheter is passed from the internal jugular vein in the neck through the venous system in the chest and into the abdomen, where selective catheterization of the veins that drain the liver (hepatic veins) is performed. The catheter is wedged into the hepatic vein and CO_2 is injected into the catheter. The CO_2 diffuses through the abnormal liver tissue and into the portal vein, providing a visual target for shunt placement on subtracted angiographic images (Figure 2.1). A needle is then passed through the liver from the hepatic vein to the portal vein, and once the portal vein is successfully accessed, a guidewire is advanced into the portal vein. A metallic stent is deployed and expanded along the length of the liver tissue track, creating a conduit for blood flow from the portal vein to the hepatic vein. A successful TIPS procedure will decrease the blood pressure in the portal vein, relieving the symptoms described above.

The single most difficult step in TIPS creation is accessing the portal vein [31, 86]. CO_2 portography, a method of imaging the portal vein by injecting carbon dioxide into the easily accessible hepatic vein, is often sufficient to provide guidance for TIPS placement; however, up to 10 needle passes through the liver may be required to access the portal vein [78]. In difficult cases, direct ultrasound-guided puncture of the portal vein through the abdominal wall

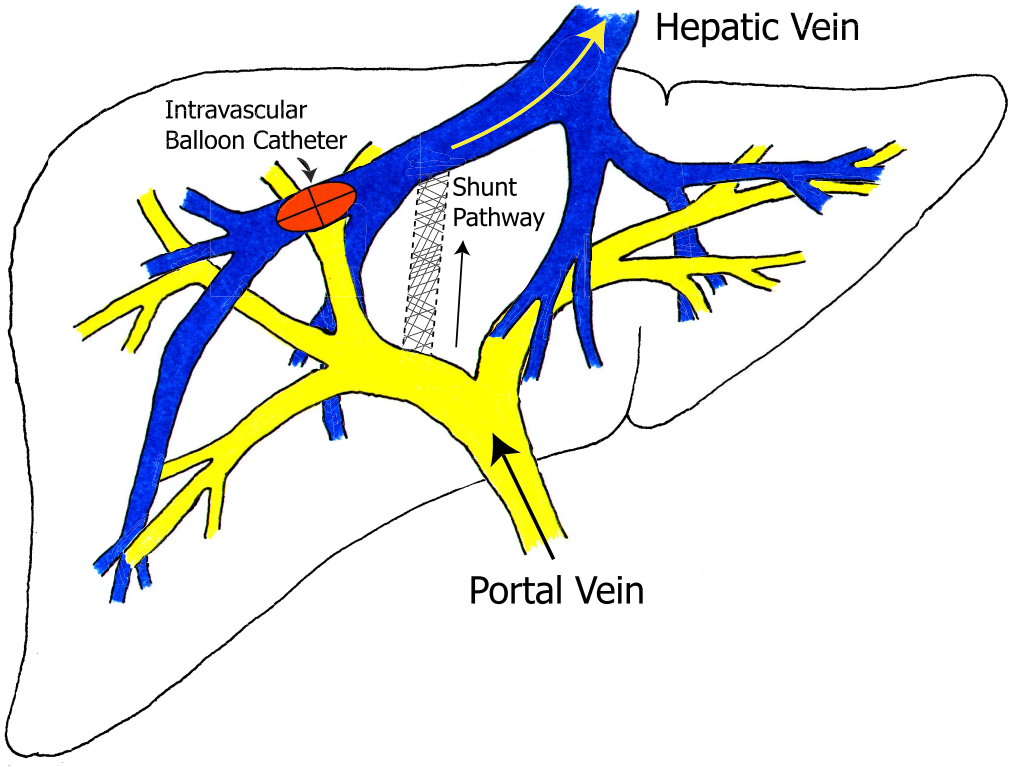


Figure 2.2: Relationships of the vessels in the liver and the desired TIPS shunt location. The liver is viewed from the front, as if sliced parallel to the plane of the body (coronal section). The shunt should connect the portal venous system (yellow) to the hepatic venous outflow (blue).

may be necessary to adequately locate the portal vein prior to shunt formation [33, 78]. More recent modifications in image guidance used during TIPS creation include the use of biplane fluoroscopy or three-dimensional ultrasound (US), both of which aid in locating the portal venous system in the anterior-posterior plane [10, 79]. However, the use of ultrasound requires transducer manipulation at a location distant from where the interventionalist is manipulating the catheter. The technique is difficult, cumbersome, and requires two operators. Even with the use of this burdensome augmentation, an average of 4 needle passes is required [79]. Because of the transparency of the liver and its blood vessels to x-rays, biplane fluoroscopy fails to provide information about displacement of the liver (and the target portal vein) due to breathing motion and direct force from the needle, so the interventionalist is aiming for a moving target whose location is largely unknown. Complication rates resulting from TIPS procedures are variable, but have been reported in 10-30% of patients [6]. Most serious complications result from needle mispunctures during attempted portal vein access. Significant mispunctures have been reported in up to 30% of some series [31] and may result in death. While complication rates have improved from earlier studies, complications and high radiation doses do remain challenges in this procedure [65, 66].

2.2 Existing Guidance Systems and Related Research

Several groups have been pursuing research on accurate intraoperative guidance. Jolesz and colleagues at Harvard created a neurosurgical guidance system using a 'double doughnut' operative magnetic resonance (MR) imaging system (GE Medical Systems) and a flat display; this system provides 2D overlays of anatomical and surgical features [8, 46, 87, 90]. Hawkes, King and associates at King's College have published extensively on 3D registration for surgical guidance systems, particularly for applications in the liver [24, 52, 63, 73, 74]. Sauer, Bani-Hashemi, and colleagues of Siemens Corporate Research in New Jersey are developing a head-mounted display (HMD) system for medical visualization and guidance [67, 85]. Fuchs' and Pisano's research group at UNC has an augmented reality (AR) system based on a head-mounted display that has been used for human clinical trials. This system was shown to yield more accurate biopsy placement in breast phantoms than conventional ultrasound [82]. Many other commercial systems exist, particularly in the fields of orthopedics, ENT surgery, and

neurosurgery, but these are less closely related to the topic of this work and will not be directly discussed here.

Chapter 3

OVERVIEW OF THE INTEGRATED GUIDANCE SYSTEM

3.1 Design of the Proposed Guidance System

Our goal for improved TIPS guidance is a fused 3D model of the liver vessels, tissues and the needle in correct relationship to one another, updated in real-time. The steps involved are as follows:

1. Segmentation of the vasculature and liver from CT and MR to form a 3D model of the blood vessels

2. Intraoperative 3D/2D registration of the vessel model with biplane (stereo) DSA[1] images

3. Reconstruction of the needle (used to create a passage between the two venous systems) into 3D from biplane images

4. Image-based tracking of portal vein motion and needle position during the respiratory cycle and during needle pressure in 5 degrees of freedom from biplane images

5. Display of the relevant objects to the surgeon using a 3D display system (Figure 3.1)

The following sections describe each component of the planned system and its relationship to the other pieces of the system.

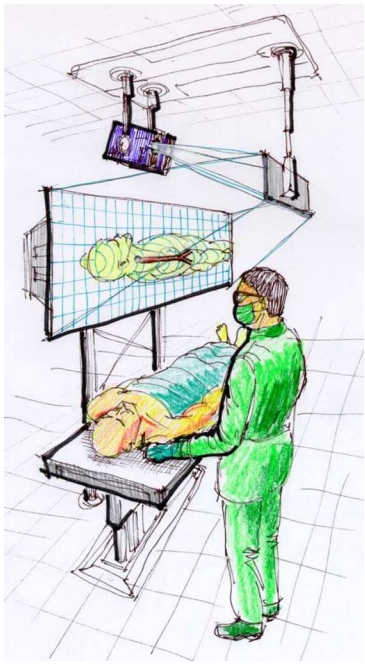


Figure 3.1: Artist's rendition of the intraoperative guidance system. Sketch by Andrei State.

3.2 Vessel Modeling

The planned system uses pre-operative diagnostic imaging to construct a high-fidelity three-dimensional model of each patient's hepatic vessel system[2] Aylward has developed and validated highly accurate and reliable methods for determining the locally cylindrical medial model that best matches a given set of pre-operative images [1–3, 12]. Chapter C.1 provides additional information. Figure 3.2 shows an example of an extracted vessel model.

These models are expressed in the coordinate system of the pre-operative imaging device. This coordinate system is referred to as the *model* coordinate system. This coordinate system will be connected to the patient and camera coordinate systems during the 2D-3D registration step described below and in Chapter 7.

[1] Digital Subtraction Angiography — An x-ray imaging technique in which a background image is subtracted from a contrast angiogram, resulting in improved contrast between the blood vessels and all other structures

[2] Preoperative imaging generally consists of contrasted CT or MRI. CT imaging is attempted at maximum inspiration with breath holding, but this is often impossible due to the patient's respiratory condition. Respiratory gating may also be attempted to improve image quality.

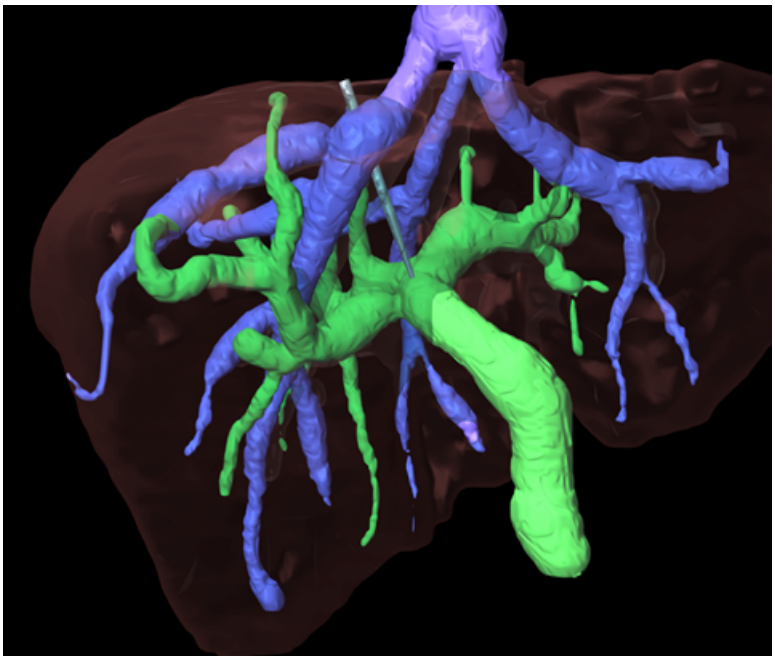


Figure 3.2: Three-dimensional model of the blood vessels in a patient's liver, extracted from CT. A synthetic needle has been added to show the usual path of the TIPS needle from hepatic to portal veins. The image was rendered using 3D Studio Max^{TM} and is viewed from the front of the patient.

3.3 Fluoroscope Calibration

The methods described in Chapter 4 are used to determine the intrinsic and extrinsic parameters of the biplane fluoroscopy system. Our system has one camera with a posterior-to-anterior (PA; back to front) view and another with an orthogonal lateral view (Figure 3.3). For ease and generality of notation, the PA camera system is referred to as camera A and the lateral system is referred to as camera B. The calibration step yields $\mathbf{M}_A^{INT}$ and $\mathbf{M}_B^{INT}$, the intrinsic parameter matrices that project objects from the camera A or B coordinate systems on to the corresponding image planes, respectively. The calibration step also yields $\mathbf{R}_{B\rightarrow A}$ and $\vec{t}_{B\rightarrow A}$. This rotation matrix and translation vector are used to transform coordinates from camera B's frame to camera A's frame according to the formula

$$\vec{X}_A = \mathbf{R}_{B\rightarrow A}\vec{X}_B + \vec{t}_{B\rightarrow A}$$

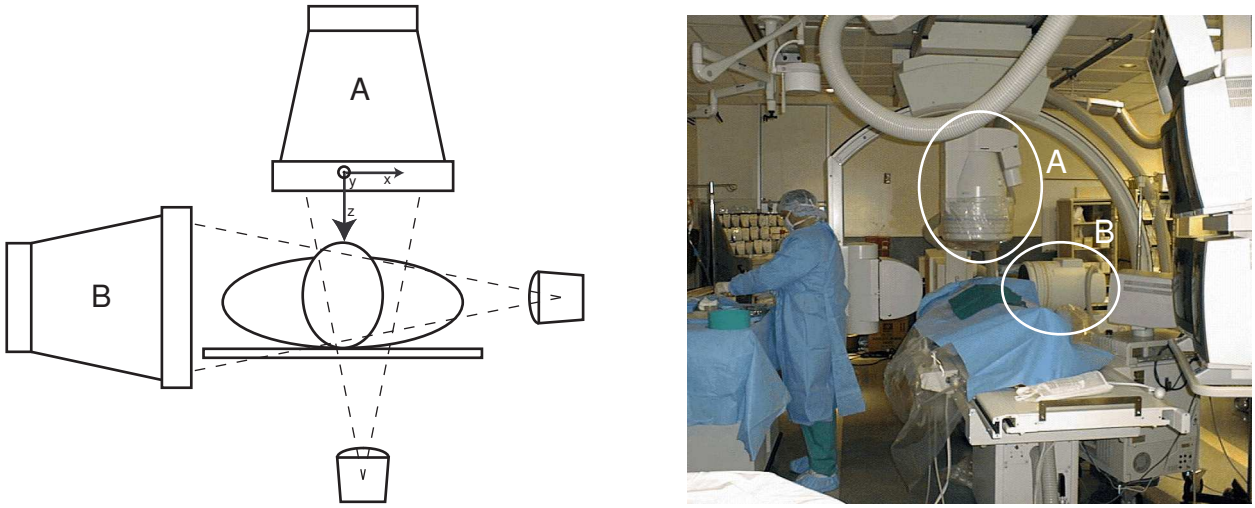


Figure 3.3: Relationship of fluoroscopic cameras to the patient. There are two separate cameras; one images from back to front, while the other images through the patient's side (left). The right image shows a biplane fluoroscopic imaging suite at UNC Hospitals.

where $\vec{X}_A$ is a 3D vector of coordinates in A's system and $\vec{X}_B$ is the corresponding 3D coordinate vector in B's system. It may be helpful to think of $\vec{t}_{B\to A}$ as the origin of camera B expressed in the coordinates of camera A and the rows of $\mathbf{R}_{B\to A}$ as the unit basis vectors of B expressed in terms of the unit basis vectors of A. This convention is used for all coordinate frames in the system.

The coordinate frame of camera A is used as the central reference frame throughout the integrated system because it does not change during the procedure and is referenced in many of the computations within the system.

3.4 Intravascular Balloon Tracking

As described in Chapter 5, a contrast-filled intravascular balloon is used as a proxy tracking fiducial for the hepatic vessels. By following the balloon's position and orientation over time, the change in position and orientation of the hepatic vessels can be inferred. Since the balloon is tracked using the biplane fluoroscopy system, the resulting measurements can be expressed as updates to the transformations linking the balloon and camera A coordinate systems, as

follows:

$$\vec{X}_A = \mathbf{R}_{Ball \to A} \vec{X}_{Ball} + \vec{t}_{Ball \to A}$$

This transformation is updated with each new pair of images. The update rate varies from a few frames per seconds to thirty frames per second depending upon the operator and current system settings.

It is important to note that the ellipsoidal balloon can only be tracked in three degrees of position and two degrees of orientation because it is symmetric about its long axis. The rotation about that axis is assumed to be static after the initial measurement. See page 68 for additional discussion of this issue.

3.5 Vessel 2D-3D Registration

This system uses 2D-3D registration to establish the relationship between the preoperative 3D vessel model and the biplane 2D imaging system. A novel approach to this problem is described in Chapter 7. The result of this registration is the transformation from model coordinates to camera A coordinates (3D) at the start of imaging, as follows:

$$\vec{X}_A = \mathbf{R}_{model \to A} \vec{X}_{model} + \vec{t}_{model \to A}$$

This is combined with the balloon tracking results from the first frame of the tracking series to link the model and balloon coordinate systems, as follows:

$$\begin{aligned} \vec{X}_{balloon} &= \mathbf{R}_{A \to balloon} \left(\mathbf{R}_{model \to A} \vec{X}_{model} + \vec{t}_{model \to A} \right) + \vec{t}_{A \to balloon} \\ \vec{X}_{balloon} &= \mathbf{R}_{model \to balloon} \vec{X}_{model} + \vec{t}_{model \to balloon} \end{aligned}$$

This transformation is then used to relate the preoperative model to the intraoperative coordinate system of camera A via the real-time balloon-to-camera transformation updates. The final transformation from model coordinates to camera A coordinates is as follows:

$$\vec{X}_A = \mathbf{R}_{balloon \to A} \left(\mathbf{R}_{model \to balloon} \vec{X}_{model} + \vec{t}_{model \to balloon} \right) + \vec{t}_{balloon \to A}$$

3.6 Needle Tracking

During the TIPS procedure, a large-bore (14-18g), rigid needle is passed through the vena cava, into the hepatic vein, and pushed through the hepatic tissue and into the portal vein. In this system, the needle will be tracked in real-time by extracting it from the intraoperative fluoroscopic images and reconstructing the visible components into three dimensions. Because the needle is reconstructed in the coordinate frame of the camera A system, there is no need to translate between coordinate systems. The updated needle reconstruction coordinates will be fed into the overall intraoperative model after each tracking iteration.

3.7 Intraoperative Display

The final coordinate frame for this system is that of the intraoperative display. The original vision of this display system was a stereo-projection display with a tracked user, but other approaches are under exploration. Regardless of the final display used, the display fundamentally acts as a window into the virtual world. The view through this window can be thought of as a two-part transformation: the first part places the user's eyes in the virtual world, and the second part places a display surface somewhere along the path between the user and the objects of interest. The specifics of the display system and its transformations will be material for a future work.

3.8 Hardware Platform

The current version of the system operates on two synchronized PCs on an isolated network (Dell . Each system carries a Foresight Imaging I-50 video capture card capable of digitizing images up to 2048x2048 at 12-bit grayscale depth. The system is connected to two analog video output ports on a Siemens Neurostar biplane fluoroscopy system in the UNC Department of Radiology. The images appear to be anisotropically compressed from the native digital images before they are output to the analog ports, so an unequal aspect ratio is noticeable in the visible figures (and in the calibration results). Balloon and needle tracking will occur concurrently on each machine, and the results will be merged together to update the master model. Preliminary results in our tracking work suggest that this design will be sufficient to

handle the processing of the two image streams at an interactive rate (up to 30Hz).

Part II

Calibration of Biplane Fluoroscopes

The second part of this dissertation will explore a novel method for accurately and efficiently calibrating a biplane fluoroscopy system. The proposed method combines bundle adjustment techniques with fluoroscopic distortion correction using a single calibration target. The section closes with a discussion of the accuracy of reconstruction with this method versus that of traditional approaches.

Chapter 4

CALIBRATION OF FLUOROSCOPIC IMAGING SYSTEM

Fluoroscopy, or real-time x-ray imaging, is commonly used to look 'inside' a patient during medical procedures. The images acquired by such a system provide a *projective* view of the scene, meaning that features are sampled by a bundle of rays radiating outward from the focal point or, equivalently, appear larger or smaller according to their depth from the camera. A ray passing from the focal point of a camera and through a feature in an image must eventually intersect the object represented by that feature, but the depth of that object along the ray is undefined by the ray alone. If one wants to make accurate three-dimensional measurements of the object, it is necessary to bring in additional information, such as focus, texture, lighting, or, in the case of this thesis, additional views of the scene.

Two unique views of an object feature define, in the general case, two distinct rays in three-dimensional space (one for the feature in each camera). The intersection of these rays occurs at the three-dimensional location of that feature, and the intersection of rays for many such features on an object define the surface of that object. In order to perform such intersections, however, it is necessary to explicitly or implicitly determine the geometry characteristics of each camera as well as the spatial relationship between them. A thorough determination of the imaging geometry will allow measurements in the three-dimensional coordinate space of the world using only the two-dimensional feature locations in each image.

As discussed below, there are many options for gathering and representing the geometry of the imaging system. I chose to use explicit calibration with a custom-manufactured calibration object because this approach allows highly accurate object reconstructions [41, 102, 103]. In order to achieve this goal, I first developed an x-ray compatible calibration phantom with known feature locations (§4.2.1). I then extended modern computer vision techniques like

bundle adjustment to accommodate the specific challenges of the fluoroscopic imaging system (§4.2.4). A validation study of these methods (§4.4) showed that the global, multi-frame bundle adjustment method with distortion correction yielded a mean three-dimensional reconstruction error of 0.438mm in a 20x20x20cm working volume.

4.1 Background

Calibration is the process through which the parameters describing a system become known. In the case of fluoroscopy or machine vision, a successful calibration process yields an accurate representation of the imaging system's geometry. This representation allows one to track, reconstruct, and otherwise interpret the three-dimensional world from a set of two-dimensional projective views.

The body of scientific literature related to the calibration of fluoroscopic systems can be broadly divided into two sections: works in the realm of machine vision, most commonly published in journals such as IEEE-PAMI or the International Journal of Computer Vision; and works published specifically on medical fluoroscopy, which most often appear in journals such as Medical Physics. Perhaps because they tend to be published in journals with different audiences, the works in these two sections are not always well connected and have developed in somewhat different manners. The machine vision literature has included extensive developments on quantitative methods for single-view calibration [102, 103], epipolar geometry [27], and higher-order camera configurations [41, 100]. Pollefeys et al have developed methods for scene reconstruction using uncalibration sequences, possibly involving multiple cameras [76, 77]. The fluoroscopy literature has had some developments in geometric calibration methods [29, 64], but has had much more effort dedicated to correction of the large distortion errors associated with fluoroscopic imaging [35].

Fluoroscopy and optical imaging have several important differences that must be considered when bridging knowledge between the systems. Most objects are partially transparent when viewed with x-ray fluoroscopy, which means that correspondence ambiguities between views can not be solved using the simple occlusion cues that generally exist in optical imaging. In addition, fluoroscopy systems generally have narrow fields-of-view (6-12° for Siemens Neurostar) relative to most optical imaging systems, which often have fields-of-view of 25°

or more. The narrow field-of-view makes estimation of focal length, optical center, and depth substantially more difficult by limiting the perspective effect throughout the image.

Much of the research in fluoroscopy has been conducted with single camera systems because biplane systems are both uncommon and costly. Unlike the parallel stereo setups that are common in machine vision, biplane fluoroscopy tends to involve two cameras with orthogonal views that intersect in the volume of interest. This means that calibration objects must be unambiguous and provide sufficient coverage of two simultaneous orthogonal views in order to be useful for full stereo calibration with a single pair of frames. Such a setup does, however, offer the possibility of improved determination of focal length and translation depth in each view due to the orthogonal view provided by the second camera. This work presents a novel approach to exploiting the features of an orthogonal biplane system to yield improved calibration and 3D reconstruction results.

Calibration for machine vision extends at least as far back as Sutherland's work in the 1960s [98]. Since that time, several prominent works have shaped the science of calibration. The first paper proposing a direct linear estimate of the pose between two cameras occurred in 1981 [61]. Beginning in 1986, Roger Tsai and colleagues published a series of works on methods for explicit calibration of intrinsic and extrinsic parameters using known calibration objects [102, 103]. Faugeras and colleagues published extensively on calibration, including linear methods for direct calculation of the projection matrix [26, 27]. Hartley and Triggs, among others, have developed methods for addressing higher-order camera systems [41, 100].

The fluoroscopy literature shares some similarities with that of machine vision. Starting in 1989, Fencil and Metz published several papers on the calibration of biplane fluoroscopy systems using correspondences of unknown geometry [29,64]. This paralleled the development of self-calibration techniques in the machine vision literature. Fluoroscopy, however, has also had to address the extensive distortion associated with x-ray image intensifier tubes. Early work focused on the correction of radial distortion effects [9], while later works involved more complex models of distortion [16, 35, 36, 57, 84].

4.1.1 Machine Vision Literature

The concept of measuring a user's view and augmenting it with synthetic imagery is generally credited to Sutherland's early paper on the topic [98]. In 1981, H. C. Longuet-Higgins published his seminal paper on a linear method of determined the fundamental matrix using eight or more correlated points in two views [61]. His formulation of the fundamental matrix and the equations that form the foundation of epipolar geometry have persisted to date with little change.

Direct Parameter Calculation

Tsai described an efficient approach to single-view calibration using a known calibration object. Tsai's method provides a linear solution to the intrinsic and extrinsic camera parameters, but requires a nonlinear optimization step if radial distortion is to be factored in. Tsai reported an average error of 0.4mils in a working volume 2 inches across, or about 1 part in 5000 [102,103]. Faugeras et. al. introduced alternative methods of calibrating an optical system via direct determination of the projection matrix. These works included linear methods for calculating the projection matrix from a set of matched points [26,27].

Correlated Sequence with Known Object

Borghese and Cerveri described a method of calibrating camera pairs using a simple rigid bar of known length. The positions of two fiducial markers on the ends of the bar are followed throughout a series of frames, and the resulting accumulation of correspondences is used to estimate the calibration of the cameras. [11].

Auto-calibration in an Unknown Scene

Pollefeys et al. developed methods to calibrate and reconstruct simultaneously using no prior knowledge of cameras or scene geometry [76,77]. Scene correspondences are robustly estimated from sequential frames, and then these correspondences are used to estimate camera parameters across frames. The resulting parameters are then used to check epipolar constraints on the original correspondences and to find (many) new correspondences. After some number of

iterations of this process, the refined set of camera parameters is used to densely reconstruct the unknown scene into three dimensions.

Bundle Adjustment

The term 'bundle adjustment' refers to a set of techniques for optimizing solutions to vision problems. Bundle adjustment generally includes the simultaneous optimization of view (i.e., camera) and scene (i.e., calibration target) parameters, the results of which are provably optimal. A large variety of cost functions and optimization methods exist, a few of which will be described below (§4.2.4). For additional detail, Triggs et al provide an outstanding and comprehensive survey of bundle adjustment methods [99].

Distortion in Optical Systems

The major source of distortion in optical systems is the system of lenses used to focus light onto the imaging device. If a ray from the optical system's focal point to an object is different than the ray from that object to the first element of the imaging system, then the image of that object will be distorted. This is sometimes intentional, as in the case of fish eye lenses that are designed to image a large field of view using a relatively narrow optical path. Because it is exceedingly costly to manufacture perfectly curved lenses, most optical systems also have some inaccuracy in their optical pathways. Rays passing through the center of a series of lenses generally intersect each lens perpendicularly, experiencing relatively little refraction and thus little influence of imperfect lens curvature. Rays that travel farther from the center of the optical path are refracted more strongly by each lens in the series, so their paths through the system are influenced more heavily by imperfectly shaped lenses. This effect leads to distortion that is dependent upon a ray's radius from the center of the optical system, which is briefly termed radial distortion. Convex lenses that are too flat cause peripheral rays to appear farther from the center than they should (so-called pincushion distortion), whereas convex lenses that are too curved cause such rays to appear closer to the optical center than they should (known as barrel distortion).

4.1.2 Fluoroscopy Literature

Geometric Calibration

The fluoroscopy literature has generally emphasized the determination of rigid transformations relating views (extrinsic imaging parameters) over measurements of the intrinsic camera parameters. Fencil & Metz published a series of prominent works on the determination of the extrinsic transformations of biplane imaging systems using unstructured, corresponding scene features [17, 29, 64]. In their original paper, they adapted Longuet-Higgins' and Tsai's methods to accommodate the inherent redundancy of biplane views. Their method allows calibration of extrinsic parameters and reconstruction of scene points with no formal phantom. Instead, at least eight corresponding points are identified in the two views and used to determine the extrinsic transformation matrix. They do not discuss methods of determining the intrinsic imaging parameters. Simulations in their subsequent work suggested that their methods would yield 1-2mm of RMS error [29].

In 1998, Esthappan and colleagues published another work on the determination of the extrinsic parameters of a biplane imaging system /citeesthappan98. Their method relied upon a calibration phantom of known dimensions (a Lucite cube embedded with 1mm spherical lead targets) as a base for establishing correspondence and scale. Their phantom studies found mean errors of 2mm in their 3D reconstructions.

In Alan Liu's dissertation [59], he presented methods for calibration of the focal length and principal point of a monoplane fluoroscopic imaging system. For the calculation of focal length, he proposed that a regular, planar grid of steel wires, placed perpendicularly to the optical axis, be imaged at two different positions along that axis. He then correlated the intersection points of the wires in each image and uses the geometry of similar triangles to estimate the focal length of the system [59, pp 181-3]. In order to calibrate the principal point of the system, he proposes that a tall cylinder with two precision crosshairs, orthogonal to but separated along the cylinder's long axis, be positioned along the estimated optical axis and translated until the crosshairs overlap exactly [59, pp 183-4]. When the cylinder has been positioned correctly, the projected point of overlap will coincide with the principal point in the imaging plane. The accuracy of these methods has not been published.

Distortion Correction

As mentioned above, one of the ways in which the calibration of fluoroscopic imaging systems is different from that of optical imaging systems is that the distortion of image features follows a much more complex model in x-ray image intensifiers. Efforts to correct fluoroscopic distortion go at least as far back as the work of Bonard et al from 1978 [9]. Their original work addressed the measurement of areas on distorted images for cardiac diagnostic studies. They used the intersections of a grid of metallic wires to correlate true and distorted point coordinates. They then fit a full tenth-order bivariate polynomial function to map between correct and distorted parameters. This approach reappears in the later work on this topic.

An alternative approach to distortion correction appeared in 1981 from Wallace and Johnson [105]. Their work on orthopedic imaging required improved quantification accuracy, so they similarly proposed to use a regular grid of wires for distortion correction. Rather than providing a global solution, however, they elected to perform linear correction within each calibration grid box. The four corners of each grid box were identified, and then the trapezoidal form of each distorted box was corrected back to be a perfect square. Internal features were corrected essentially via interpolation.

In 1987, Chakraborty proposed a third approach to fluoroscopic distortion correction. This new approach explicitly modeled the projection of x-rays on to a curved input phosphor and the reprojection from the phosphor to the imaging device as two distinct processes. This method also utilized a regular wire grid to establish ground truth. While elegant, the complexity of this approach may have limited its widespread adoption.

Rudin, Bednarek, and Wong used a variant of the polynomial approach to distortion correction in their 1991 paper [84]. They elected to focus on the distortion induced by the curved input phosphors of the time, so they chose to use only the third- and fifth-order radial distance terms for their correction. They used a hexagonal grid of precision holes with spherical bearings to establish ground truth. They found that the use of both terms was superior to the use of a single term, but they did not compare their method to higher-order polynomial models. During this year, another group used a full third-order bivariate polynomial approach to build a real-time distortion correction display system [37].

Söderman and colleagues wanted to correct distortion in stereotactic digital subtraction angiograms. This group also used a regular wire mesh placed over the image intensifier, and they used an approach similar to that of Wallace and Johnson. Distortion offsets were calculated for each intersection viewed on the grid, and bilinear interpolation was used to refine the correction for points between the intersections. [91]

Fahrig et al were concerned with correcting image intensifier distortion in a rotating x-ray gantry to improve their computed tomographic reconstructions [25]. They mounted a regular grid of beads to their image intensifier to serve as known fiducials; fifth-order bivariate polynomials were fit for each angular view using a least-squares approach to correct for any distortion. This approach resulted in mean residual errors of about 0.07 pixels.

Gronenschild has published several substantial papers on fluoroscopic distortion correction [35, 36]. He emphasizes the 'global' approach to distortion correction in which a single set of equations models the distortion for the full image, similar to Bonard's 1978 work. He used a regular wire grid to establish ground truth. He then tested full bivariate polynomials from third to ninth order on several different fluoroscopes. He found that fifth order polynomials were sufficient for correction on most fluoroscopes, though some required as high as seventh order for optimal results. His approach yielded mean errors of about 0.16 pixels. His later paper compared local interpolation methods to his global method and found that the global method offered four to tenfold improvement in correction accuracy.

Ruijie Liu, Rudin, and Bednarek [60] extended Gronenschild's 1997 work to accommodate the position of a rotating C-arm fluoroscope. They used a hexagonal grid of bearings to establish truth (see Rudin et al above) and a polynomial correction model. They found that fitting full fifth-order bivariate polynomials to the coordinates at ten-degree intervals and then fitting polynomial functions of the angular position to the correction coefficients yielded the best results (0.20 pixels).

Alan Liu's dissertation discusses several additional works on this topic [59].

4.2 Methods

4.2.1 Calibration Target for Fluoroscopy

Based upon the principles described in the previous section, an ideal calibration target for fluoroscopic systems should have the following characteristics:

- Non-coplanar correspondences - The set of 2D-3D correspondences must include at least six points in each view, no four of which are co-planar.

- Radio-opaque features - Since fluoroscopy uses high-energy x-rays to form images, the target's features must absorb x-rays near 80keV in order to be visible.

- Discernable features on projection - Corresponding features must be discernible under projective transformations. For example, it may be impossible to distinguish between overlapping spheres under x-ray projection.

- Spans the working volume - Regions outside of the calibration object may be incorrectly modeled, so the calibration object should span the full working volume.

- Accurate manufacturing of target - It must be possible to manufacture the phantom in a reliable manner.

- Accurate localization of features in image - It must be possible to localize the key coordinates of features in the output images.

- Unique correspondence between views - There should be only one possible correspondence between the multiple projective views.

I have designed a calibration target using two orthogonal plexiglass planes and spherical metal bearings, as shown in Figure 4.1 [80]. The two planes are mounted between thick plexiglass plates with precision-milled grooves to maintain the planes' orthogonality. Each plane has a pattern of drilled holes that accept the spherical bearings. The bearings have a brass coating with a zinc core, so they are non-ferromagnetic. The phantom is compatible

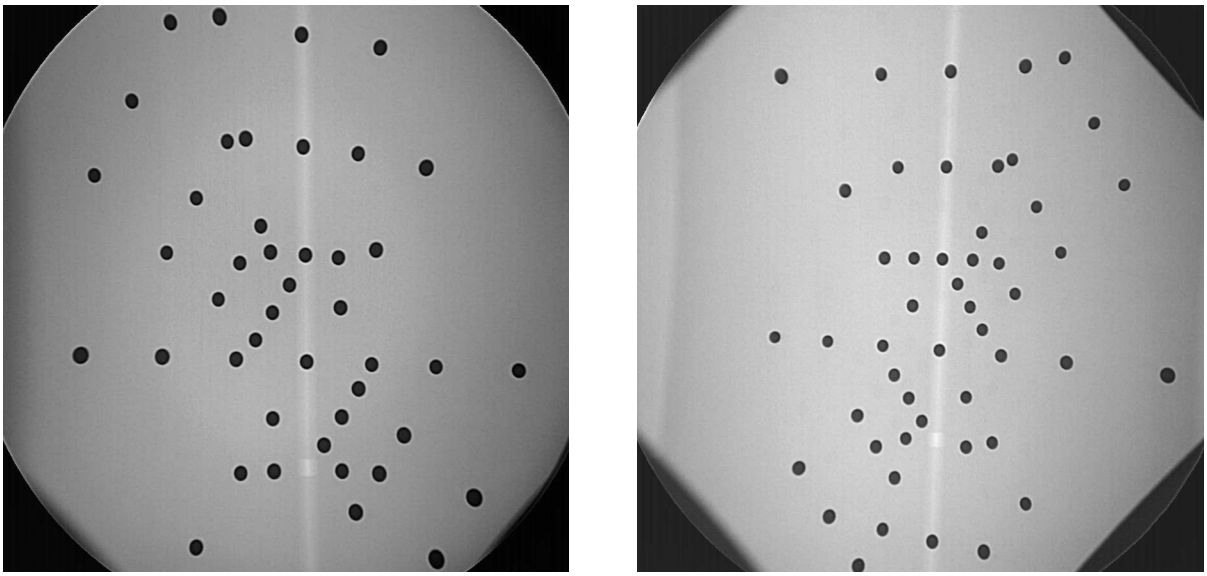


Figure 4.1: X-ray images of the fluoroscope calibration target (PA left; lateral right).

with x-ray, fluoroscopy, computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound, so it could be used to establish transformations between modalities.

4.2.2 Feature Extraction

For each image in the data set, all of the fully visible calibration features were automatically extracted by thresholding the image and identifying connected regions with empirically appropriate size and roundness characteristics. A threshold was selected using Otsu's method (Matlab's `graythresh` function), which minimizes the within-class variance of the binary populations [69]. The resulting binary image features were filled to eliminate unwanted holes and labelled using eight-neighbor connected components analysis. Each binary object was analyzed for appropriate area, eccentricity, and roundness (area:perimeter ratio) properties to filter out irrelevant features. The location of each target was calculated as the grayscale centroid of its elliptical region[1].

[1] Interestingly, this measurement is not quite correct – the centroid of the elliptical projection of a sphere does not project back to the true center of the sphere unless that center lies along the optical axis of the camera. This issue will be dealt with briefly in the future work section of this dissertation.

4.2.3 Calibration of a Single View

The simplest form of calibration is the estimation of parameters from a single frame of a single camera view. Methods for single-frame calibration of one or two views have been well-described by Sutherland, Longuet-Higgins, Tsai, Faugeras, and Hartley among others [26, 27, 42, 61, 98, 102, 103]. In this work, I used the normalized direct linear transformation (NDLT) method with a non-linear optimization step to estimate the parameters of each frame and view, as described in Hartley's text on the subject and summarized below [42].

Correspondence between the three-dimensional model points and their two- dimensional projections was estimated manually for the first frame in a series and automatically for each subsequent frame in a series. These feature locations and their corresponding ideal model locations were used as input to the NDLT method to estimate the parameters for each frame and view.

The NDLT method calculates the maximum likelihood estimate of the projective transformation by minimizing the two-dimensional reprojection errors of the known calibration points. The method includes four steps as follows:

1. Normalization - The image and space coordinates are normalized to improve the stability of the subsequent numerical methods. The 2D image points are transformed so that their centroid lies at the origin and their RMS distance from the origin equals $\sqrt{2}$. The 3D space points are similarly transformed so that their centroid lies at the origin and their RMS distance from the origin equals $\sqrt{3}$. The transformed points are defined as $\tilde{x}_i = Tx_i$ and $\tilde{X}_i = UX_i$ for the 2D and 3D points respectively.

2. Linear transformation - The system of equations relating the 2D and 3D points $\tilde{x}_i = \tilde{P}\tilde{X}_i$ is rearranged and stacked into a $2n$ x 12 matrix A such that $Ap = 0$ for noiseless data. We arbitrarily set $\|p\| = 1$ to constrain the scale of the solution. The optimal solution to p in $Ap = 0$ is the singular vector associated with the smallest singular value of A.

3. Nonlinear optimization - The next step is a nonlinear minimization of the residual 2D reprojection error $\sum_i d\left(\tilde{x}_i, \tilde{P}\tilde{X}_i\right)^2$. The linear estimate of $\tilde{P}$ is used as the initial state

of the projection matrix. This implementation uses the Levenberg-Marquardt variant of Newton's method.

4. Denormalization - The final estimate of $\tilde{P}$ is transformed back to the original coordinate systems using the transformation $P = T^{-1}\tilde{P}U$.

This implementation was validated using synthetic data and was found to produce parameters correct to a factor of approximately 1e-10.

4.2.4 Simultaneous Two-View, Multi-Frame Calibration

When calibrating a camera based upon a single frame, errors in target point localization or in the calibration phantom can yield inaccurate estimates of the parameters for that frame. As the view varies across frames, these errors can yield variations in the calibration results that appear as jumps between frames - a highly undesirable feature of any calibration method (see Figure 4.2 and 4.3). In the interest of finding an optimum calibration across such a series of frames, I developed a nonlinear optimization scheme based on bundle adjustment that simultaneously optimizes the global camera calibration parameters and the local frame transformations. This approach also allows the incorporation of distortion correction and model correction methods into the calibration process.

First, we assume that the true intrinsic parameters for both cameras and the true intercamera extrinsic parameters are constant across the series of frames. We then define K_A and K_B as the 3x3 intrinsic projection matrices for the A and B (anteroposterior and lateral) cameras, respectively. Each of these upper triangular matrices has five degrees of freedom (four if skew is assumed to be zero, as in this work) and is defined up to a scale factor. We also define $\mathbf{R}_{A\to B}$ and $\vec{t}_{A\to B}$ as the intercamera extrinsic rotation matrix and translation vector, respectively (a 3D Euclidean transform). These describe the transformation $\vec{X}_B = \mathbf{R}_{A\to B}\vec{X}_A + \vec{t}_{A\to B}$ and add another six degrees of freedom to the system.

If the calibration phantom and/or cameras are moved during the sequence, then the extrinsic transformation mapping the world or calibration object coordinate system to that of the camera system will vary by frame. We define $\mathbf{R}^i_{W\to A}$ and $\vec{t}^i_{W\to A}$ as the rotation matrices and translation vectors comprising the Euclidean transformation from World to camera A

coordinates for each frame i. The per-frame extrinsic estimates add $6N$ parameters to the system, where N is the number of pairs of frames in the sequence. The transformations are composed in the same manner as for the intercamera transformation above.

This joint optimization process comprises three major steps: convergence of the shared parameters across the views and frames without model or distortion correction, joint optimization using a single set of shared parameters for all of the frames, and a final optimization step that uses shared parameters and optional model and distortion correction. The model and distortion correction methods were excluded from the convergence and joint optimization steps because inclusion of the additional parameters in these initial steps was found to yield inferior results; the optimizer appeared to settle into local minima created by the additional parameters. Each of these steps is described below.

Interframe Parameter Convergence Process

The first step in this method seeks to converge the parameter estimates for each frame to a common solution. The method begins with intrinsic and extrinsic parameter estimates for each camera in each frame in the series as determined above. These parameters are optimized using a cost function that incorporates the variances of the shared parameters as cost elements.

The least-squares function to be optimized, $Cost(\mathbf{x})$, is described below. This cost function has two components: $Err(\mathbf{x})$, the sum of the geometric error distances between each image point and the projection of the corresponding model point; and $Var(\mathbf{x})$, the weighted sum of the variances of the parameters that are being globalized. In these equations, m represents the number of frame pairs, n represents the number of model points, $P_i^{A|B}$ represents the projection matrix for camera A or B, X_j^{mod} represents the three-dimensional coordinates of the j^{th} model point, and $x_j^{A|B}$ represents the 2D image coordinate of the j^{th} on camera A or B. In addition, Γ represents the set of parameters to be globalized, as described above, p_{li} represents the l^{th} such parameter in the i^{th} frame, and μ_l is the mean of the l^{th} parameters across all frames.

$$Cost(\mathbf{x}) = Err(\mathbf{x}) + Var(\mathbf{x}) \tag{4.1}$$

$$Err(\mathbf{x}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(P_i^A X_j^{mod} - x_j^A\right)^2 + \left(P_i^B X_j^{mod} - x_j^B\right)^2 \tag{4.2}$$

$$Var(\mathbf{x}) = \sum_{l\in\Gamma} w_l \frac{\sum_{i=1}^{m}\left(p_{li} - \mu_l\right)^2}{m-1} \tag{4.3}$$

The variance weighting factor w_l was chosen as the inverse of the initial variance of the corresponding parameter in order to equalize the 'pressure' across parameters with differing magnitudes. In addition, this weighting factor was multiplied by a scale factor increasing from 0.25 (2^{-2}) to 256 (2^8) by even powers of two over 6 sequential iterations of the optimizer; this increasing cost is intended to allow a gradual convergence from local parameterization to global parameterization.

The bundle adjustment approach minimizes the sum of the squared cost elements via a first order linear optimization. This optimization can be considered in terms of Newton's method, though the final implementation is more complex. The optimization process requires the calculation of the first derivative matrix of the cost function (a.k.a. the Jacobian) and the subsequent calculation of the optimal next step, as follows:

$$Cost(\mathbf{x_{t+1}}) = Cost(\mathbf{x_t}) + \frac{\partial Cost(\mathbf{x})}{\partial \mathbf{x}}\Delta\mathbf{x_t} + \epsilon = 0 \tag{4.4}$$

$$\Delta\mathbf{x_t} = -\left(\frac{\partial Cost(\mathbf{x_t})}{\partial \mathbf{x_t}}\right)^{-1} Cost(\mathbf{x_t}) \tag{4.5}$$

$$\Delta\mathbf{x_t} = -J^{-1}(\mathbf{x_t})Cost(\mathbf{x_t}) \tag{4.6}$$

$$\mathbf{x_{t+1}} = \mathbf{x_t} + \Delta\mathbf{x_t} \tag{4.7}$$

The global optimization routine uses a region trust method (Matlab's `lsqnonlin` function for large scale problems) with a preconditioned conjugate gradient (PCG) step [19, 20]. The PCG step essentially rescales the elements of the function's Jacobian matrix so that it is well-conditioned and then finds a 'good' subspace for optimization[2]. The region trust method adapts the optimization steps to better fit the relative scale and accuracy of the cost function element derivative approximations. These methods are critical to achieving convergence

[2] This is a gross simplification of these methods; please refer to the citations above for a more thorough and accurate description

when the partial derivatives of some parameters are significantly different than those of other parameters.

The matrix of partial derivatives, the Jacobian matrix, is explicitly calculated in this implementation. While the Jacobian may be approximated by finite differencing, this requires one calculation per element of the matrix for each iteration. For large problems (600x1000 or more), such an approach is impractical for regular use. Instead, the Jacobian is explicitly calculated at each step using the partial derivative equations shown in Appendix A.

Joint Optimization with Shared Parameters

After the bundle adjustment process described above converges, we are left with a set of per-frame parameters with substantially reduced variance. As a subsequent step in this process, it is desirable to reformulate the parameterization so that the parameters that should be shared across frames are treated as such in the optimization. To this goal, the shared parameters are estimated as the means of the corresponding per-frame parameters from the prior optimization step. We then have 4 intrinsic parameters for each camera and 6 intercamera extrinsic parameters, for a total of 14 shared parameters. The world to camera frame extrinsic parameters remain distinct for each frame due to motion of the world coordinate system, which is bound to the calibration object, relative to the camera frame. These add another $6m$ parameters to the system. The cost function is also updated to eliminate the variance cost elements, as follows:

$$Cost(\mathbf{x}) = Err(\mathbf{x}) \tag{4.8}$$

$$Err(\mathbf{x}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(P_i^A X_j^{mod} - x_j^A\right)^2 + \left(P_i^B X_j^{mod} - x_j^B\right)^2 \tag{4.9}$$

The output of this process is the optimal shared-parameter calibration for this data set. The presence of other errors in the system, particularly calibration model errors and fluoroscopic image distortion, suggests that a final optimization step is necessary.

Calibration Model Correction

As will be described in section 4.4, static errors in the target model became apparent in the analysis of residual errors. Manufacturing errors are a highly plausible explanation for these difficulties. Prior work in bundle adjustment has included model estimation or correction as an integral component of the optimization [99]. Uncertainty in the calibration model points can be accommodated by adding global parameters describing the true location of each model point. A set of coefficients $\kappa_j \in \Re^1$ defining the j^{th} bearing's displacement along the x-, y-, and z-axes are used to compensate for the uncertain position along these drill axes. The corrected model coordinates are given by the equation:

$$\hat{X}_j^{mod} = \begin{bmatrix} x_j^{mod} + \kappa_{x_j} \\ y_j^{mod} + \kappa_{y_j} \\ z_j^{mod} + \kappa_{z_j} \end{bmatrix} \tag{4.10}$$

Integrated Distortion Correction

Gronenschild et al found that the optimal approach to distortion correction in x-ray fluoroscopy was to model that distortion using a full fifth degree polynomial in each image coordinate [35, 36]. This approach was added as an option to the optimizers described above. When distortion correction is enabled, the measured image coordinates are adjusted using the following equations:

$$\hat{x} = k_{20}^x x^2 + k_{11}^x xy + k_{02}^x y^2 + k_{30}^x x^3 + k_{21}^x x^2 y + \cdots + k_{14}^x xy^4 + k_{05}^x y^5 \tag{4.11}$$

$$\hat{y} = k_{20}^y x^2 + k_{11}^y xy + k_{02}^y y^2 + k_{30}^y x^3 + k_{21}^y x^2 y + \cdots + k_{14}^y xy^4 + k_{05}^y y^5 \tag{4.12}$$

The fixed and first-order terms are excluded because of their redundancy with the optical center and pixel scale parameters. The distortion correction coefficients $k_a^{x|y}b$ are added to the parameterization described above.

In order to successfully implement this method, it was necessary to temporarily rescale the pixel coordinates used in the distortion correction calculations. Pixel coordinates in the range of 1000 became computationally unmanageable when raised to the fifth power; very small changes in the high-order coefficients yielded dramatic swings in corrected coordinates that prevented the optimizer from converging efficiently. To avoid this problem, pixel coordinates

were scaled into the range [1.5, 2.5] within the distortion correction calculations. This approach yielded manageable magnitudes for the high-order terms and allowed the optimizer to converge expediently.

Final Optimization Step

The final step in this optimization incorporates model and/or distortion correction into the shared parameter optimization approach described above. The cost function used for this step is as follows:

$$Cost(\mathbf{x}) = Err(\mathbf{x}) + Constr(\mathbf{x}) \tag{4.13}$$

$$Err(\mathbf{x}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(P_i^A X_j^{mod} - x_j^A\right)^2 + \left(P_i^B X_j^{mod} - x_j^B\right)^2 \tag{4.14}$$

$$Constr(\mathbf{x}) = m\sum_{j=1}^{n}\left(\kappa_{x_j}^2 + \kappa_{y_j}^2 + \kappa_{z_j}^2\right) \tag{4.15}$$

This cost function has two components: $Err(\mathbf{x})$, the sum of the geometric error distances between each image point and the projection of the corresponding model point; and $Constr(\mathbf{x})$, a cost element intended to restrict the correction of the model points. The output of this step is the best estimate for the shared parameters of the system, the structure of the calibration model, and the distortion in the system.

Summary of Complete Method

The final bundle adjustment algorithm consists of four discrete steps, as follows:

1. Calibration of each individual frame using the normalized DLT method

2. Optimization of the redundant per-frame parameterization with minimization of reprojection errors and progressive weighting on the variance of shared parameters

3. Reduction of the parameterization to use a common set of intrinsic and intercamera extrinsics parameters across all frames with subsequent optimization of the reprojection errors in the reduced parameter space

4. Final optimization step incorporating model and distortion correction; additional cost elements are added based on the magnitude of the model correction vectors

The output of this final step includes the optimal parameterization of the shared camera intrinsics and intercamera extrinsics, the fifth-order distortion correction parameters for each camera, and the corrections to the 'ideal' calibration model.

4.3 Design of Validation Studies

The methods described above were evaluated by capturing a large series of images of the calibration phantom, calibrating the system using subsets of those images, and then reconstructing the phantom's target locations from a separate subset of images. This approach allowed comparison of reconstruction accuracy between calibration methods, numbers of frames used in each calibration, and the presence of model and distortion correction in each calibration. The detailed methods are described in the following sections.

4.3.1 Data Collection and Allocation

Image Capture

A custom imaging phantom containing radio-opaque metal bearings in two orthogonal planes was used to provide radiographically distinct reference targets for calibration. Several series of biplane views of the calibration phantom were captured in a single session, yielding a complete data set comprising 509 pairs of images. In order to vary the calibration images over the session, the phantom was translated and rotated during and between these imaging series. The biplane fluoroscopes remained stationary throughout the session. The images were captured from an 884x884x12-bit grayscale analog video signal and converted into 884x884x8-bit grayscale images using Foresight Imaging I-50 video capture cards.

Image Processing and Calibration

For each image in the data set, all of the fully visible calibration targets were extracted using the methods described in §4.2.2. Correspondence between points was estimated manually for the first frame pair and automatically for all frames thereafter. The only constraint on the

manual registration is that correct correspondence must be established between a predominance of points in each view and the ideal 3D model. When sufficient correct correspondence is present, the subsequent steps will reduce registration errors and correct some errors in correspondence. The intrinsic and extrinsic parameters for each frame pair were estimated separately using the normalized direct linear transformation method and refined via nonlinear optimization of the in-plane geometric distance error. Both methods are described well in [42] and are based upon work by Sutherland, Hartley, and others. This implementation of the single-frame calibration method was validated using noiseless synthetic data and was found to produce parameters correct to a factor of approximately 1e-10 in this noiseless context.

Partitioning into Calibration and Testing Sets

The sequence of available frames was partitioned into 15 interleaved sets of 32 images each. Within each set, the first sixteen (16) frames were allocated for experimental calibration and the last sixteen (16) frames were allocated for evaluation of reconstruction accuracy. The images within the calibration and evaluation subsets were linked for evaluation purposes so that any calibration subset that included the n^{th} calibration frame had the n^{th} evaluation frame as part of its evaluation subset. The calibration subset was further partitioned into subsets with 1 (x16), 2 (x8), 4 (x4), 8 (x2), or 16 (x1) frames per subset, with overlap allowed between the subsets. Subsets were selected to maximize the time interval between frames within all subsets. This structured pairing allows for improved analysis between calibration methods.

Calibration was performed as follows:

- For subsets with only one frame, the local calibration method (normalized DLT) was performed. No model or distortion correction was performed.

- For subsets with two frames, global calibration was performed without model or distortion correction.

- For subsets with four or more frames, global calibration was performed with and without model and distortion correction.

This pattern of testing yields sixteen local calibrations, eight two-frame global calibrations (no corrections), eight four-frame calibrations (four with and four without corrections), four eight-frame calibrations (two with and two without corrections), and two sixteen-frame calibrations (one with and one without corrections). Each subset was evaluated using the corresponding order and number of frames from its linked evaluation subset.

4.3.2 Intermediate Analysis

Reconstruction of Three-Dimensional Points

In this work, three-dimensional points were reconstructed from corresponding calibrated projection images by calculating the midpoint of the shortest line segment connecting the projection rays from each image. The details of the method can be found in [101] or [42].

Registration of Reconstructed and Ideal Models

The calibration target may have been moved between the calibration and testing frames, so the world to camera transformations estimated by the calibration routine must be updated to account for possible motion. Because we have established correlation between model points in all of the frames, we can use this correlation data to estimate the transformation between the ideal and reconstructed model points. As described in [30] and based upon [28,44], a rigid transformation can be estimated using a least-squares approach with quaternions. We seek the transformation that minimizes the error function

$$\epsilon = \sum_{i=1}^{n} \left| \vec{X}_i^{model} - \left(\mathbf{R}\vec{X}_i^{recon} + \vec{t} \right) \right|^2 \tag{4.16}$$

Minima occur where

$$\frac{\delta\epsilon}{\delta\vec{t}} = 0 \quad \text{and} \quad \frac{\delta\epsilon}{\delta\mathbf{R}} = 0$$

Explicitly,

$$\frac{\delta\epsilon}{\delta\vec{t}} = -2\sum_{i=1}^{n} \left(\vec{X}_i^{model} - \left(\mathbf{R}\vec{X}_i^{recon} + \vec{t} \right) \right) \tag{4.17}$$

$$= 2n\vec{t} - 2\sum_{i=1}^{n} \vec{X}_i^{model} + 2\sum_{i=1}^{n} \mathbf{R}\vec{X}_i^{recon} \tag{4.18}$$

and thus, by substitution into (4.3.2)

$$\vec{t} = \frac{1}{n}\sum_{i=1}^{n} \vec{X}_i^{model} - \frac{1}{n}\sum_{i=1}^{n} \mathbf{R}\vec{X}_i^{recon} \tag{4.19}$$

This equation can be substituted into (4.16) and the result simplified to

$$\epsilon = \sum_{i=1}^{n} \left| \vec{X}_i^{model} - \left(\mathbf{R}\vec{X}_i^{recon} + \frac{1}{n}\sum_{j=1}^{n} \vec{X}_j^{model} - \frac{1}{n}\sum_{j=1}^{n} \mathbf{R}\vec{X}_j^{recon} \right) \right|^2 \tag{4.20}$$

$$\epsilon = \sum_{i=1}^{n} \left| \left(\vec{X}_i^{model} - \frac{1}{n}\sum_{j=1}^{n} \vec{X}_j^{model} \right) - \left(\mathbf{R}\vec{X}_i^{recon} - \frac{1}{n}\sum_{j=1}^{n} \mathbf{R}\vec{X}_j^{recon} \right) \right|^2 \tag{4.21}$$

$$\epsilon = \sum_{i=1}^{n} \left| \hat{X}_i^{model} - \mathbf{R}\hat{X}_i^{recon} \right|^2 \tag{4.22}$$

where

$$\hat{X}_i^{model} = \vec{X}_i^{model} - \frac{1}{n}\sum_{j=1}^{n} \vec{X}_j^{model} \quad \text{and} \quad \hat{X}_i^{recon} = \vec{X}_i^{recon} - \frac{1}{n}\sum_{j=1}^{n} \vec{X}_j^{recon} \tag{4.23}$$

Rather than using the linear rotation matrix $\mathbf{R}$ to model rotation in this context, one may instead use quaternions to model rotation more robustly[3]. One can re-express (4.22) using quaternions as

$$\epsilon = \sum_{i=1}^{n} \left| \hat{X}_i^{model} - q * \hat{X}_i^{recon} * q^* \right|^2 \tag{4.24}$$

where q denotes the quaternion, '$*$' denotes quaternion multiplication, and q^* denotes the complex conjugate of the quaternion. The quaternion q is composed of a scalar q_0 and a hypercomplex vector $\vec{v}$; q can also be denoted as $[q_0, \vec{v}]$ for simplicity.

In order to preserve the scale of the vectors under transformation, it is necessary to require

[3] Rotations based upon Euler angles, such as those used throughout this text, possess singularities that disrupt the smoothness of solutions near those singular points (also known as 'gimbal lock'). Quaternions provide an alternative approach to rotations that are smooth and lack singularities. See [55] for an excellent discussion of this topic.

that $|q|^2 = 1$. One can then use this constraint to simplify the expression of (4.24) as follows:

$$\begin{aligned}
\epsilon &= \sum_{i=1}^{n} \left| \hat{X}_i^{model} - q * \hat{X}_i^{recon} * q^* \right|^2 |q|^2 \\
&= \sum_{i=1}^{n} \left| \hat{X}_i^{model} * q - q * \hat{X}_i^{recon} * q^* * q \right|^2 \qquad \text{but } q^* * q = 1, \text{ so} \\
&= \sum_{i=1}^{n} \left| \hat{X}_i^{model} * q - q * \hat{X}_i^{recon} \right|^2 \\
&= \sum_{i=1}^{n} \left(\hat{X}_i^{model} * q - q * \hat{X}_i^{recon} \right)^T \left(\hat{X}_i^{model} * q - q * \hat{X}_i^{recon} \right)
\end{aligned}$$

Considering only the repeated term and carrying out the quaternion multiplication, one finds:

$$\begin{aligned}
\left(\hat{X}_i^{model} * q - q * \hat{X}_i^{recon} \right) &= \left(\begin{bmatrix} 0 \\ \hat{X}_i^{model} \end{bmatrix} * \begin{bmatrix} q_0 \\ \vec{v} \end{bmatrix} - \begin{bmatrix} q_0 \\ \vec{v} \end{bmatrix} * \begin{bmatrix} 0 \\ \hat{X}_i^{recon} \end{bmatrix} \right) \\
&= \left(\begin{bmatrix} -\hat{X}_i^{model^T} \cdot \vec{v} \\ q_0 \, \hat{X}_i^{model} + \hat{X}_i^{model} \times \vec{v} \end{bmatrix} - \begin{bmatrix} -\vec{v}^T \cdot \hat{X}_i^{recon} \\ q_0 \hat{X}_i^{recon} + \vec{v} \times \hat{X}_i^{recon} \end{bmatrix} \right) \\
&= \begin{bmatrix} -\vec{v}^T \cdot \left(\hat{X}_i^{model} - \hat{X}_i^{recon} \right) \\ q_0 \left(\hat{X}_i^{model} - \hat{X}_i^{recon} \right) + \left(\hat{X}_i^{model} + \hat{X}_i^{recon} \right) \times \vec{v} \end{bmatrix}
\end{aligned}$$

The cross product $\vec{a} \times \vec{b}$ can be written in matrix multiplication form as:

$$\vec{a} \times \vec{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \vec{b} = [\vec{a}]_\times \, \vec{b} \tag{4.25}$$

Using this result, the term expansion above can be completed as:

$$\begin{aligned}
A_i = \left(\hat{X}_i^{model} * q - q * \hat{X}_i^{recon} \right) = \\
\left(\begin{bmatrix} 0 & -\left(\hat{X}_i^{model} - \hat{X}_i^{recon} \right)^T \\ \left(\hat{X}_i^{model} - \hat{X}_i^{recon} \right) & \left[\hat{X}_i^{model} + \hat{X}_i^{recon} \right]_\times \end{bmatrix} \begin{bmatrix} q_0 \\ \vec{v} \end{bmatrix} \right)
\end{aligned} \tag{4.26}$$

thus

$$\begin{aligned}
\epsilon &= \sum_{i=1}^{n} q^T A_i^T A_i q \quad \text{or} \\
\epsilon &= q^T \hat{A}^T \hat{A} q \quad \text{where} \\
\hat{A} &= \sum_{i=1}^{n} A_i
\end{aligned} \tag{4.27}$$

One can then restate this problem as a general eigenvector problem as follows:

$$\epsilon q = \hat{A}^T \hat{A} q \tag{4.28}$$

The quaternion that minimizes ϵ can thus be found as the eigenvector q corresponding to the minimum eigenvalue ϵ of $\hat{A}^T \hat{A}$.

The three-dimensional distances and error vectors between the reconstructed and ideal (or corrected) model points were calculated and used for the subsequent performance analysis.

4.3.3 Statistical Analysis of Results

The accuracy of each method was compared using the three-dimensional errors of reconstruction as previously described. For each calibration and evaluation subset above, the corresponding three-dimensional error distance was calculated for all included images and visible feature points[4]. This process created approximately 45 error distances per evaluation image per calibration method. Box-Cox correction was used to transform these data into Gaussian distributions; the power transformation $D^{0.4}$ was found to yield the best likelihood of a Gaussian fit for the data. Root-mean-scale values were also calculated within the subsets for use in error analysis.

An overall α of 0.05 was chosen for the analysis; Bonferroni correction was used for the subtest α values, which were scaled to $\alpha = 0.05/3 = 0.0166$ for each subtest. For each outcome test, a separate multivariate / repeated measures ANOVA was performed using the SAS function MIXED with the individual subset results describing repeated measures within each set (REPEATED statement). Linear model parameters included number of frames and inclusion of correction methods, as appropriate.

The first statistical test seeks to determine the significance of frame number and distortion/model correction within those calibration framesets that could accommodate both (4 or more frames). A multivariate repeated measures ANOVA was performed using frame number, distortion correction, and the interaction of frame number of distortion correction as effects under comparison in a within factorial -+*4,8,or 16 design. $\log_2$ of the number of frames was used to improve the spacing for determination of frame effect trends.

[4] The framework and some of the text of this design were created by Keith Muller, Ph.D. The analysis was conducted by the author.

The second statistical test compared the accuracy of reconstruction in non-corrected subsets across all levels of frame number. Another multivariate repeated measures ANOVA was performed using frame number as the sole effect. The logarithm of frame number was again used to improve spacing for trend determination.

4.4 Results

4.4.1 Stability and Convergence of Parameter Estimates

Extrinsic Parameters

Figure 4.2 shows the variation of the extrinsic parameters of the imaging system over a typical frame series. The extrinsic camera-to-camera parameters are expected to be stable over the image series because the cameras are stationary. The world to camera extrinsics are expected to vary because they are measured relative to the calibration phantom's internal coordinate frame and the target was moved throughout the imaging series. The per-frame calibration approach demonstrated substantial variability in the intercamera extrinsics, which is not desirable. The globally optimized parameters reflect consistent values across the series and demonstrate reasonable variation in the world coordinate transformations.

One of the challenges in calibrating fluoroscopic imaging systems is that their long focal lengths and narrow fields of view cause some of the projection parameters to be strongly covariant. If one considers the projection of some object point X_{obj} to an image point X_{im}, then this transformation is governed by the equation:

$$x_{im} = \frac{f}{z_{world} + z_{obj}} \left(x_{world} + x_{obj} \right) + oc_x \tag{4.29}$$

where x_{world} and z_{world} are the x and z components of an object-to-camera translation vector (assuming an axis-aligned object coordinate system to eliminate rotation), oc_x is the optical center offset in x, and f is the camera focal length. The variables x_{obj} and x_{world} determine the relative position of a point within the object and the object within the world, respectively. A similar equation applies for the y_{im} image coordinate.

Consider the effect of scaling the focal length and object depth by a factor k. We then

have:

$$x_{im} = \frac{kf}{kz_{world} + z_{obj}} (x_{obj} + x_{world}) + oc_x \tag{4.30}$$

$$x_{im} = \frac{kf\,(kz_{world} - z_{obj})}{(kz_{world})^2 - z_{obj}^2} (x_{obj} + x_{world}) + oc_x \tag{4.31}$$

$$x_{im} = \left(\frac{kfkz_{world}}{(kz_{world})^2 - z_{obj}^2} - \frac{kfz_{obj}}{(kz_{world})^2 - z_{obj}^2} \right) (x_{obj} + x_{world}) + oc_x \tag{4.32}$$

If $z_{obj} << kz_{world}$, then these equations reduce to:

$$x_{im} = \left(\frac{f}{z_{world}} - \frac{kfz_{obj}}{(kz_{world})^2} \right) (x_{obj} + x_{world}) + oc_x \quad \text{or even} \tag{4.33}$$

$$x_{im} = \frac{f}{z_{world}} (x_{world} + x_{obj}) + oc_x \tag{4.34}$$

If we consider the full family of object points, $\hat{x}_{obj} \in \chi_{obj}$, then we have:

$$\hat{x}_{im} = \frac{f}{z_{world}} (x_{world} + \hat{x}_{obj}) + oc_x \tag{4.35}$$

If we define κ as:

$$\kappa = \frac{f}{z_{world}} \qquad \text{then} \tag{4.36}$$

$$\hat{x}_{im} = \kappa (x_{world} + \hat{x}_{obj}) + oc_x \tag{4.37}$$

for all $\hat{x}_{obj} \in \chi_{obj}$ and for all values of f and z_{world} along the line of possible solutions $f = \kappa z_{world}$.

The last equation means that *when the object size is much smaller than the object-to-camera distance, joint scalings of the focal length and object-to-camera distance become undetectable in the image.* In the extreme case, this is equivalent to saying that the image becomes an orthographic projection of the scene; the values of the focal length and object-to-camera distance transformation are no longer unique, but instead encompass a family of solutions along the line described above. This also means that the parameter space for these linked parameters becomes flat under these conditions, which makes determination of the true optimum difficult to detect. In practical terms, measurement and system errors can cause large variations in the linked values of these parameters, making their true values difficult to determine.

The optical center and in-plane (x and y) translation parameters are also covariant. Beginning from equation 4.29 and again assuming that $\hat{z}_{obj} << z_{world}$, one can derive the relationship between these parameters as follows:

$$\hat{x}_{im} = \frac{f}{z_{world}} (x_{world} + \hat{x}_{obj}) + oc_x \tag{4.38}$$

$$\hat{x}_{im} = \frac{f}{z_{world}} \hat{x}_{obj} + \frac{f}{z_{world}} x_{world} + oc_x \tag{4.39}$$

One can then define a constant κ that reveals the covariance relationship between these parameters:

$$\kappa = \frac{f}{z_{world}} x_{world} + oc_x \qquad \text{and, by substitution,} \tag{4.40}$$

$$\hat{x}_{im} = \frac{f}{z_{world}} \hat{x}_{obj} + \kappa \tag{4.41}$$

Therefore the image coordinates of all of the object coordinates remain unchanged for all selections of oc_x and x_{world} such that

$$oc_x = \kappa - \frac{f}{z_{world}} x_{world} \tag{4.42}$$

The same principle holds true for oc_y and y_{world}.

As is shown in Figures 4.2 and 4.3, abrupt variations in the calibration parameters do appear to follow the covariance patterns described above. These abrupt variations drive the need for the parameter convergence steps described above. In future work, it may be possible to explicitly model these covariance effects after the single-frame calibration phase to improve the estimation of the true parameter values. This approach could be used to circumvent the convergence process and improve the expediency of the bundle adjustment method.

Intrinsic Parameters

Figure 4.3 shows the variation of the intrinsic (projection) parameters over the same frame series. The intrinsic parameters are expected to be stable over the image series because the cameras were not altered or zoomed during the image acquisition. As predicted in the prior section, the per-frame calibration approach demonstrated substantial variability in the intrinsic parameters of both cameras. The covariance concerns described above are believed to be responsible for most of this variation.

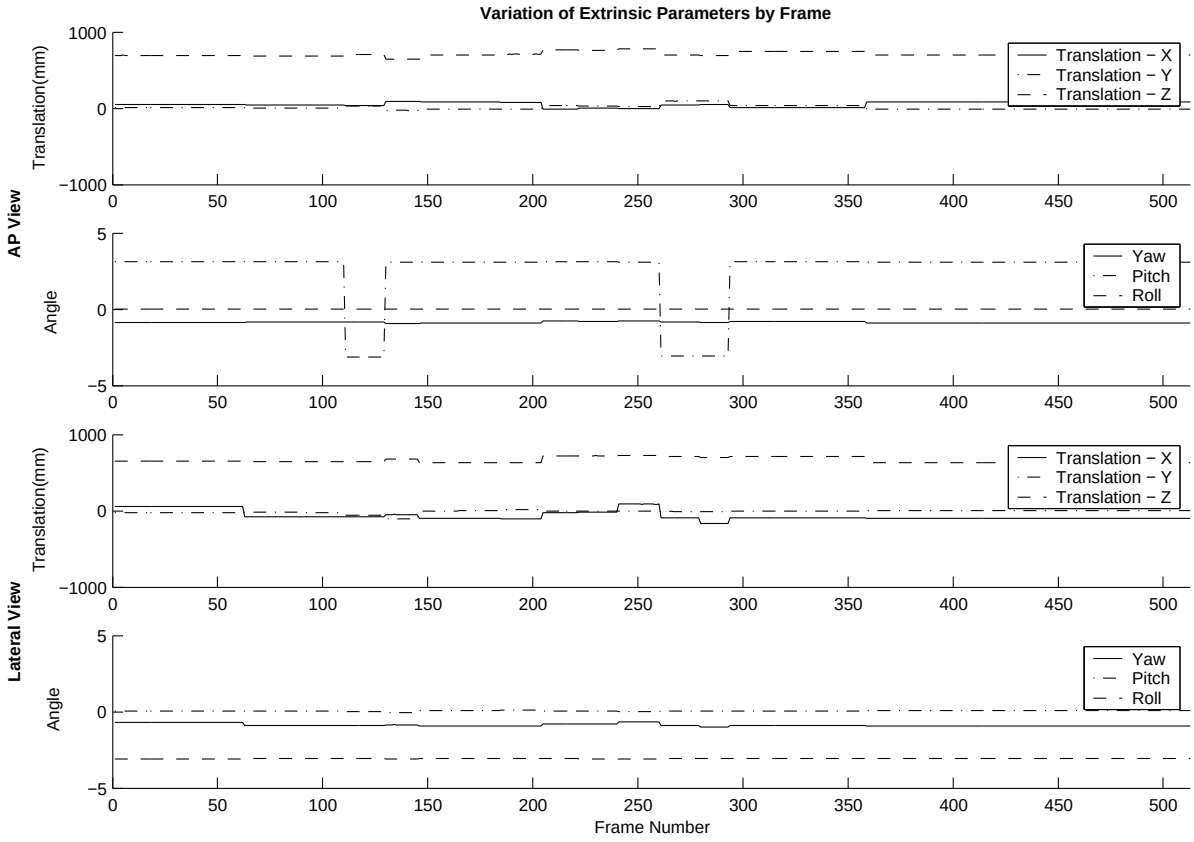


Figure 4.2: Variability of local extrinsic camera parameters across frames. Parameters were calculated for each frame and view using the methods described above. The large jumps in the angle measures on the PA view chart are due to transitions from $+\pi$ to $-\pi$, which represents a minor variation in real terms.

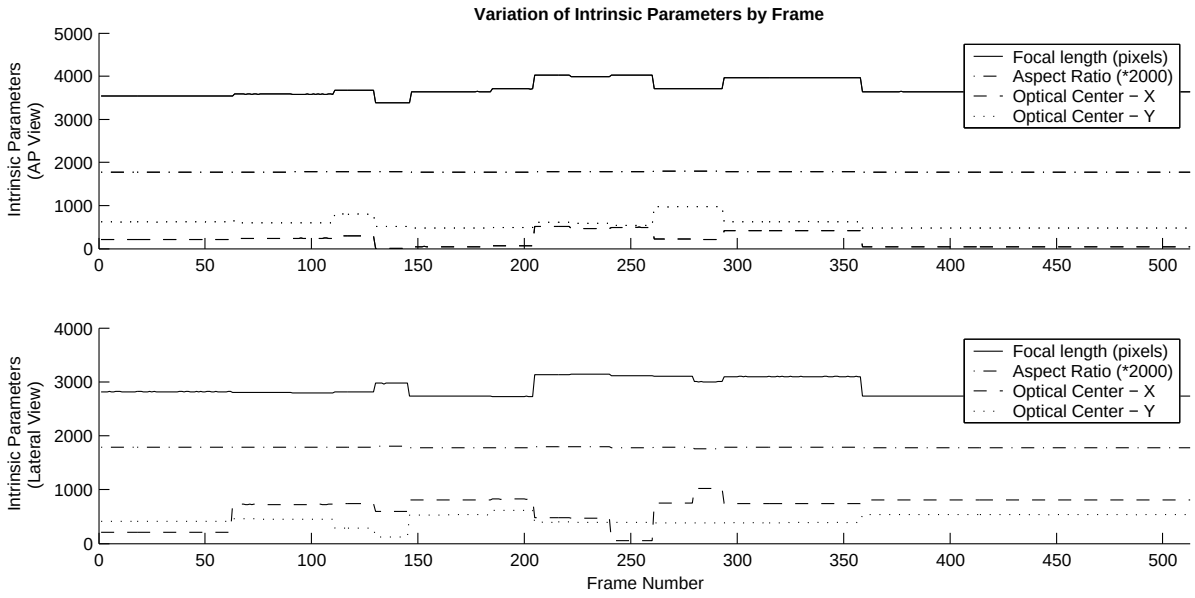


Figure 4.3: Variability of local intrinsic camera parameters across frames. Parameters were calculated for each frame and view using the approach described above.

4.4.2 Accuracy of Feature Reconstruction

The results of the reconstruction tests are reported in Figure 4.4 and Table 4.1. The mean three-dimensional distance between a 3D point reconstructed from two views and its corresponding model point was found to be 0.80mm for the per-pair calibration and 0.44mm for the globally optimized calibration[5]. The global method showed somewhat higher error variance. Figure 4.4 shows the distributions of geometric errors for each method.

[5] Distances were measured relative to the corrected model coordinates, as described in the methods section. The uncorrected model showed higher errors (+0.05mm) for both methods

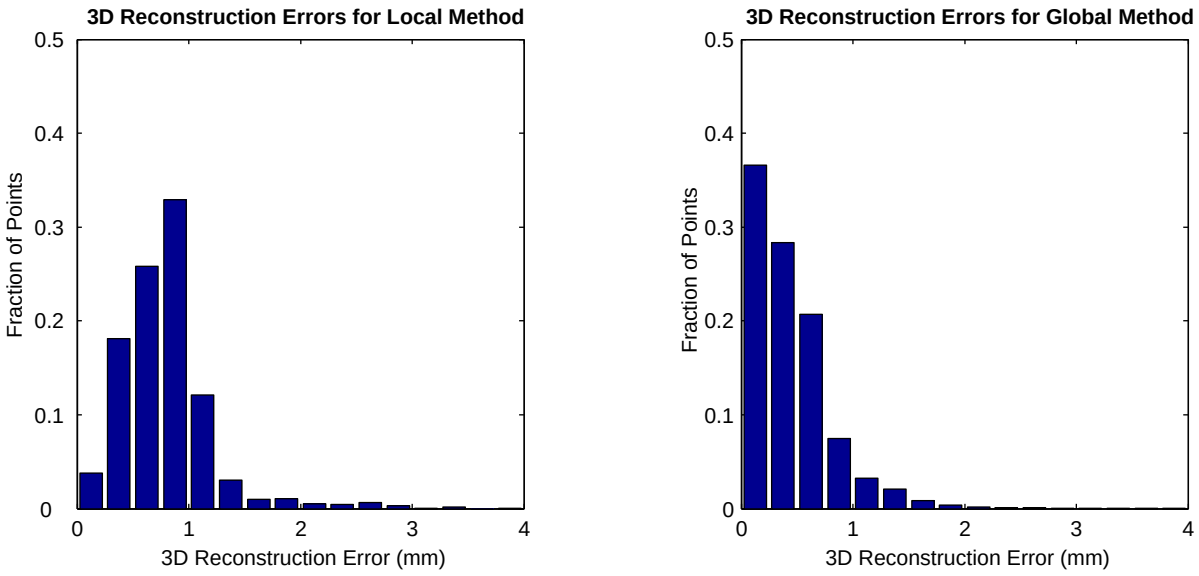


Figure 4.4: Distribution of three-dimensional geometric errors between ideal and reconstructed world points. These errors were calculated by calculating calibration parameters using one set of frames and using those results to reconstruct the phantom features in a second set of frames.

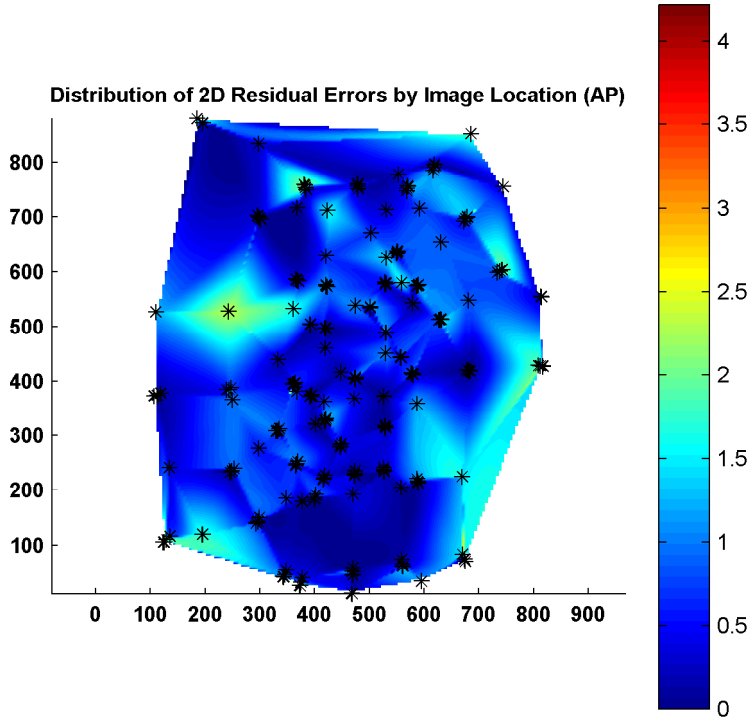


Figure 4.5: Distribution of two-dimensional residual errors on the PA image plane.

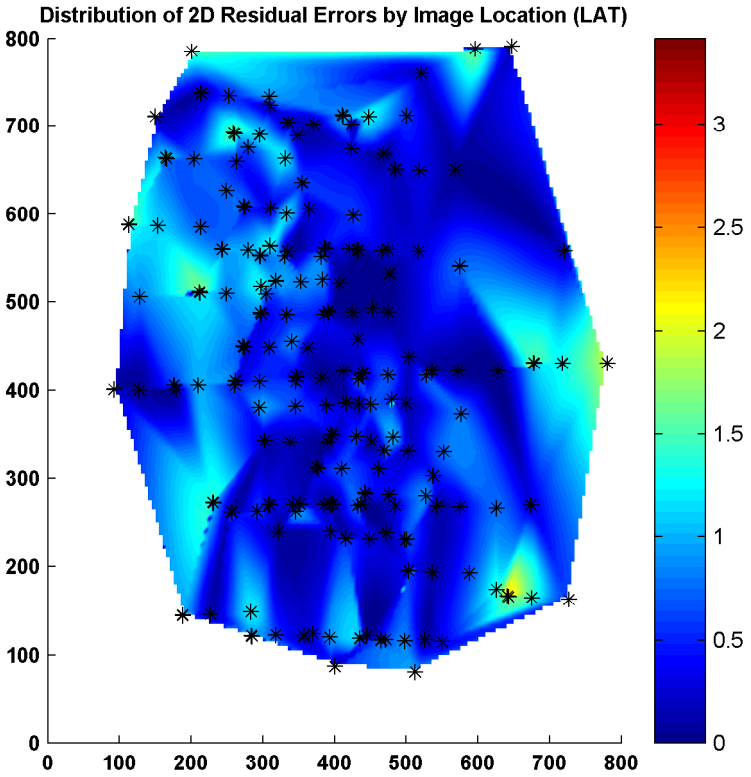


Figure 4.6: Distribution of two-dimensional residual errors on the lateral image plane.

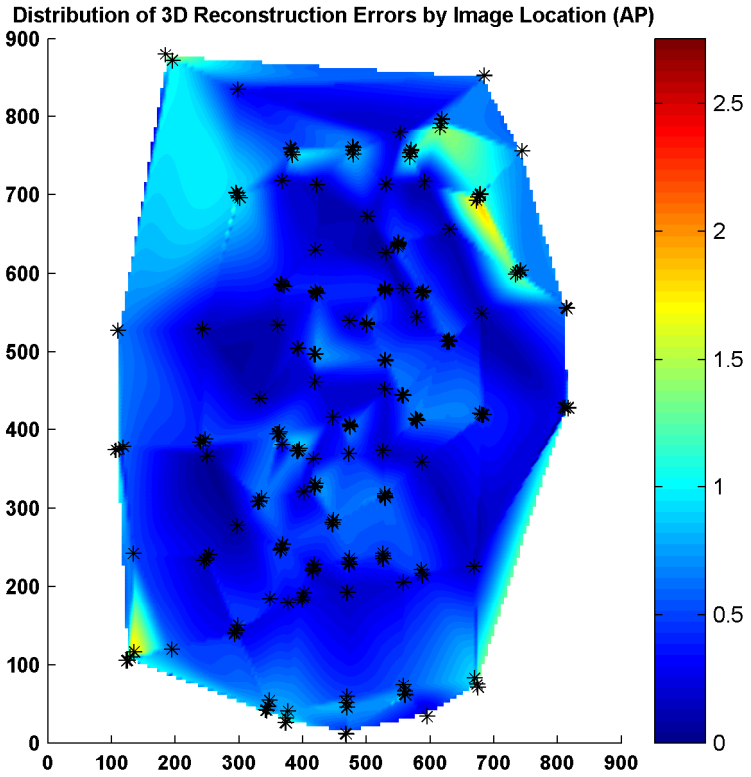


Figure 4.7: Distribution of three-dimensional geometric errors on the PA image plane.

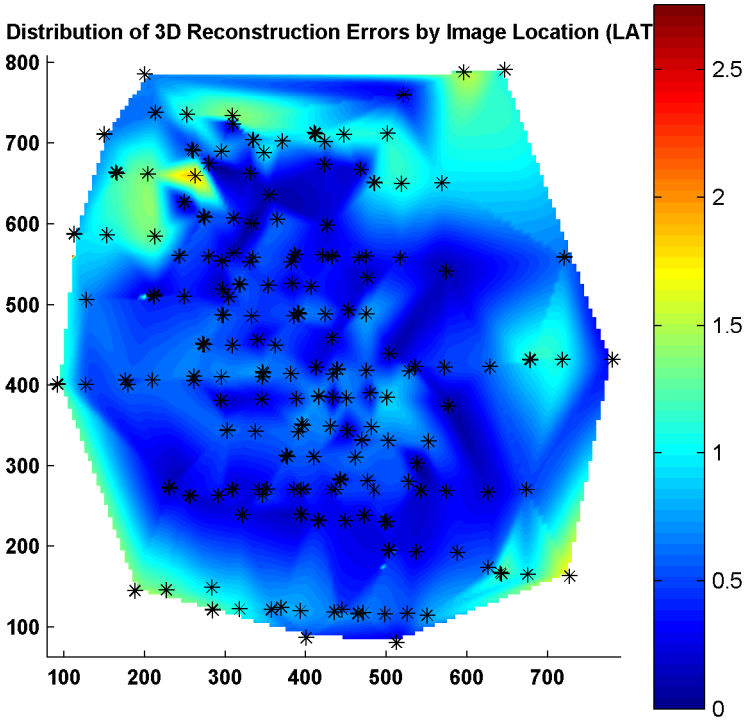


Figure 4.8: Distribution of three-dimensional geometric errors on the lateral image plane.

Table 4.1: Three-Dimensional Reconstruction Errors

Number of Frame Pairs	Without Corrections		With Corrections	
	Mean ± Std Dev	Mean(RMS) ± Std Dev	Mean ± Std Dev	Mean(RMS) ± Std Dev
1	0.796 ± 0.004	0.895 ± 0.008	.	.
2	0.804 ± 0.007	0.891 ± 0.007	.	.
4	0.807 ± 0.004	0.888 ± 0.004	0.460 ± 0.064	0.585 ± 0.120
8	0.803 ± 0.003	0.883 ± 0.004	0.440 ± 0.072	0.555 ± 0.107
16	0.803 ± 0.003	0.882 ± 0.004	0.438 ± 0.073	0.546 ± 0.095

Effects of Frame Count

The number of frames included in the optimization was found to have small but statistically significant effects upon the final reconstruction error. When distortion and model correction were not performed, analysis of the root mean square (RMS) errors shows that increasing frame counts were associated with a small but consistent decrease in those errors. Paradoxically, the use of multi-frame bundle adjustment was associated with a tiny (0.007mm) but statistically significant *increase* in the mean reconstruction error ($p < 0.001$) in this context[6]. When distortion and model correction were included, increasing the number of frames was associated with a small but statistically significant reduction in both the mean and RMS reconstruction errors ($p < 0.001$).

[6] Because RMS error was the target of optimization, it is possible to increase the mean error while decreasing the RMS error. This is generally caused by decreases in larger RMS errors that are countered by increases in smaller errors. This yields a net decrease in RMS error but a net increase in mean error

Effects of Model and Distortion Correction

The inclusion of model and distortion correction in the bundle adjustment algorithm was associated with a large and statistically significant reduction in mean reconstruction error (0.80 vs 0.44m, 0.365mm (45%) reduction, $p < 0.0001$). Figure 4.11 shows the change in per-axis error distributions with and without model and distortion correction.

Figure 4.9 shows the distributions of 3D reconstruction errors separated by calibration model point. Notice that some of the model points were associated with large or highly variable errors. These errors are believed to be due to problems with the production of the phantom. Model correction was instituted to adapt for this problem; Figure 4.10 shows the improved per-point reconstruction errors after model correction.

Distortion correction also showed a small but statistically significant independent improvement in mean errors (0.03mm reduction, $p < 0.001$). The relatively small effect of distortion correction may be due to the exceptional quality (and low distortion) of the imaging chain on this biplane fluoroscopy system, so other systems may experience more substantial effects. This result is confounded, however, by the substantial systematic errors in the uncorrected calibration model used to independently test distortion. It is unclear if distortion correction would have been more significant in the context of a highly accurate model with correspondingly less systematic positional error interfering with the detection of distortion.

4.5 Conclusions

The global calibration method was found to produce lower mean geometric errors as compared to local, per-frame calibration (0.44mm vs. 0.80mm, $p < 0.0001$). This compares favorably with the best reported mean errors of 1-2mm in the fluoroscopy literature [29, 64], and should improve the feasibility of having less than 5mm total mean error for the entire TIPS guidance chain. Correction of the ideal calibration model played a key role in this improvement by reducing the significance of systematic errors in feature localization, as demonstrated by the reduction in the magnitude and variance of errors from Figure 4.9 to Figure 4.10. Distortion correction also yielded a small (0.03mm) but statistically significant improvement in mean errors; the relatively small impact of distortion correction on this system is likely due to

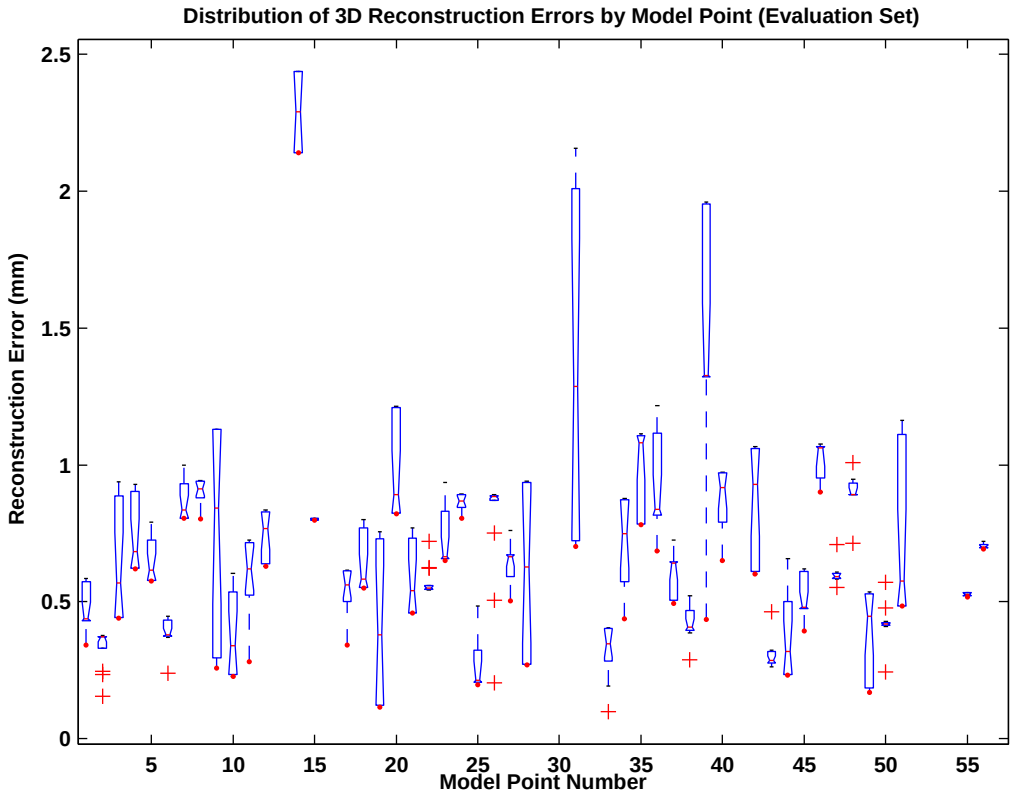


Figure 4.9: Reconstruction errors by calibration model point without model correction. Distortion correction was performed. The number and magnitude of outliers are substantial.

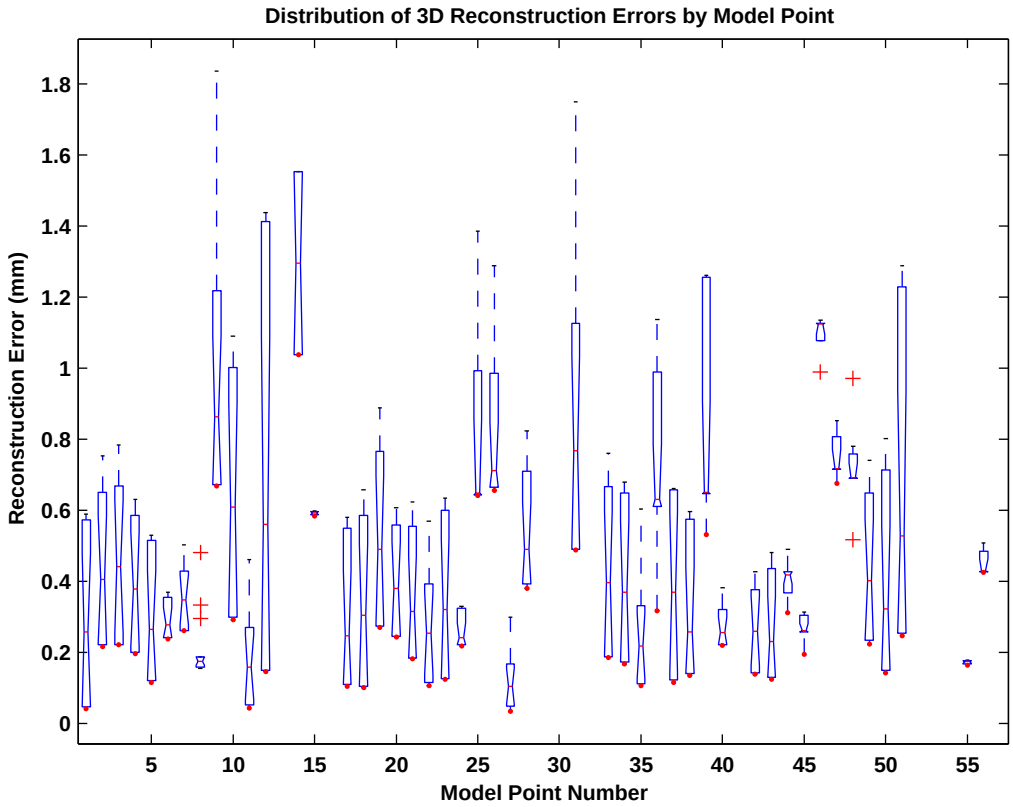


Figure 4.10: Reconstruction errors by calibration model point after model and distortion correction. The number and magnitude of outliers have been reduced.

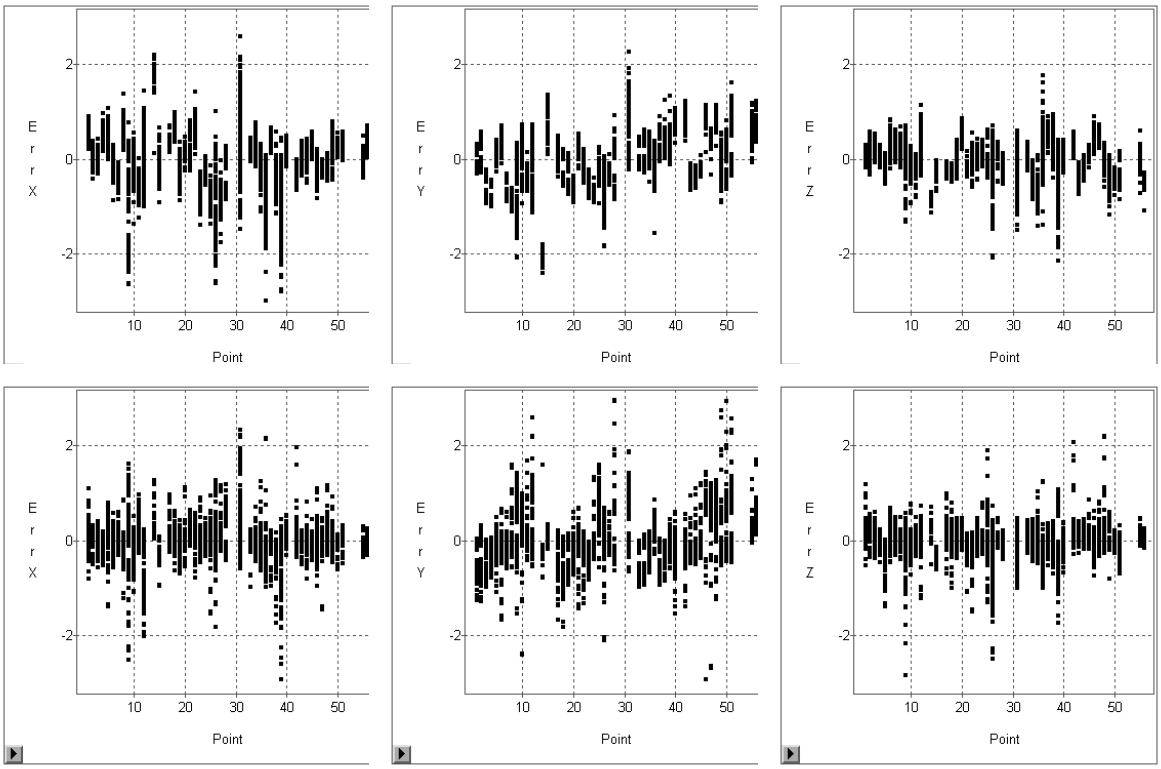


Figure 4.11: Distribution of three-dimensional geometric errors by axis, with (bottom) and without (top) model and distortion correction.

the exceptional quality of the Neurostar imaging chain, so the impact on systems with more distortion is likely to be more substantial. It is likely that the residual errors include variations in target feature localization, target manufacturing errors, and imaging system noise.

This work has shown that optimization with model and distortion correction over multiple frames can yield improvements in the reconstruction accuracy of a biplane fluoroscopy system. Future work will include improvements in image feature localization and improved target manufacturing techniques. Such improvements in reconstruction accuracy will support the demanding needs of intraoperative fluoroscopic guidance tasks.

Part III

Tracking

In this part of the dissertation, I will explore the use of fluoroscopy to track targets during a medical procedure. I will first present methods for tracking an ellipsoidal intravascular balloon in three dimensions of position and two of orientation. In Chapter 6, the integrated system will be used to characterize the motion of the human liver during the TIPS procedure using real intraoperative data.

Chapter 5

INTRAOPERATIVE ORGAN TRACKING

It is often the case that a structure of interest, such as an organ or vessel, is only visible for a short span in a medical procedure. In many procedures, the target of an intervention is not actually visible at the critical step of the intervention, so the operating physician must extrapolate the target's position based upon other information (e.g., prior views). This principle holds true in the TIPS procedure: while the target of the procedure, the portal vein, is visible during a venogram at the start of the procedure, it is completely transparent for the remainder of the procedure, including the critical needle pass from the hepatic vein to the portal vein. In the interim, the vessels are displaced by breathing motion and needle pressure and thus have unknown locations. This uncertainty is undesirable in a complex medical procedure.

When it is infeasible to directly visualize the target vessels, one could instead track a rigidly related object, such as the liver or an intrahepatic catheter, and use that object to estimate the portal vein's true location. By using computational methods to track a strongly related proxy fiducial, one can provide the operating physician with an improved estimate of the true location of the target structure. This chapter describes a method of tracking such structures using an intravascular balloon catheter viewed with two orthogonal fluoroscopic cameras.

In Section 5.1, I will discuss liver motion and existing approaches to image-based target tracking. I will then explore the use of an intravascular balloon catheter as a high-contrast intraoperative tracking fiducial (§5.2). Next, I will present the computational methods used to localize and follow this intraoperative target. This chapter will close with a discussion of future work (§5.4). Experimental results will be discussed in Chapter 6.

5.1 Background

5.1.1 Liver Motion

The liver is a large solid organ located on the right side of one's torso, just below the diaphragm. The liver has numerous attachments to surrounding structures that constrain its motion. The liver is bounded on three sides by the bony ribcage and superiorly by the diaphragm; it is additionally tethered by the left and right triangular ligaments and by the falciform ligament, all of which limit or preclude rotation. The problem of liver tracking is also eased in patients with end-stage liver disease because these patients have rigid, fibrotic, non-deformable livers that move en bloc. One can thus assume that all changes in the liver are the result of rigid transformations; the possibility of non-rigid bending or deformation is not considered in this work. Indeed, the standard of clinical practice is to assume that liver motion and deformation during needle pressure consists solely of translation along the rostro-caudal axis.

There are several causes of liver movement during the TIPS procedure. Due to the ligaments that bind it to the diaphragm, the liver moves with the regular ebbs and flows of respiration. Digestive actions of the stomach and intestines can also cause the liver to move. Finally, the force accompanying the use of intravascular surgical tools, such as pressure from the TIPS needle, may also cause the liver to move. Acting together, these forces introduce substantial motion and uncertainty into the liver tracking problem.

The TIPS procedure comprises three major stages of operation: vessel visualization, portal vein access, and shunt formation. During the vessel visualization phase, a small access port is placed in the right jugular vein in the neck and a catheter is threaded into the hepatic vein. A small amount of carbon dioxide is then injected into a small branch of the hepatic vein; this CO_2 diffuses through the abnormal hepatic parenchyma in TIPS patients and temporarily concentrates in the portal vein, yielding a subtle but visible enhancement of the portal veins. This initial portal venogram is used as an estimate for the position of the portal vein for the remainder of the procedure[1]. Figure 5.1 demonstrates one such portal venogram and its use as an overlay later in the procedure.

[1] The structures of the liver, including the portal vein, are not visible in standard fluoroscopic images, so contrast agents must be used to enhance structures of interest.

In the second phase of the procedure, a long, rigid needle is advanced through a protective sheath into the hepatic vein. The position and orientation of the needle is assessed relative to the targeted portal vein segment, and then the needle is pushed through the liver parenchyma and toward the portal vein. The liver is essentially invisible on fluoroscopy throughout the needle passes, so the physician must estimate the actual position of the target vessel using the prior portal venogram, experience, and intuition. This is a challenging visualization task – successful portal vein access is reported to take an average of ten needle passes per patient [78].

After the portal vein has been accessed, the TIPS needle is withdrawn and an intravascular shunt expansion balloon is placed. The balloon is inflated at high pressure to expand the shunt pathway and allow sufficient blood flow to relieve the patient's portal hypertension. A wire stent is placed in the shunt pathway to prevent the surrounding tissue from collapsing or growing into the newly formed shunt space. Figure 5.2 demonstrates a typical configuration of the liver vessels and TIPS shunt.

Figure 5.1: Fluoroscopic images of the hepatic vessels. The left image shows a CO_2 venogram of the hepatic and portal vessels. Such images are often acquired early in the procedure and then overlaid on subsequent views to provide an estimate of the portal vein location (right, synthesized). The true vessel location is offset from that shown by the overlay.

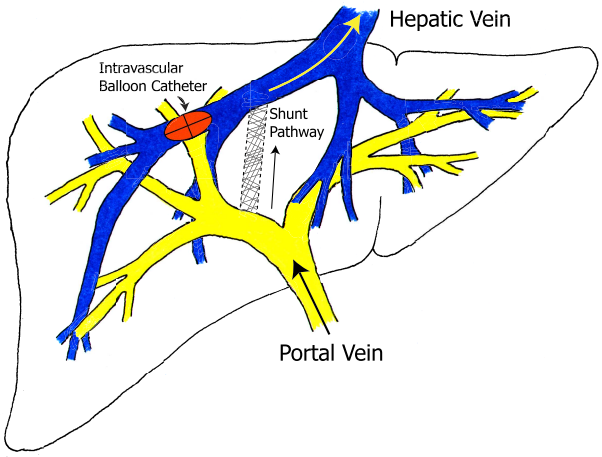


Figure 5.2: Front view of the liver, including hepatic vessels and intravascular balloon placement. The intravascular balloon is generally placed in a branch of the right hepatic vein (blue, left side of image). Arrows indicate the direction of blood flow through the vessels.

5.1.2 Current Approaches to Organ Tracking

The need to track an anatomical structure of interest is not unique to the TIPS procedure, so some insight into the problem may be gained by considering prior approaches to this problem. This section reviews existing approaches to organ tracking specifically; the next section considers target tracking as a generalized problem.

Current approaches to intraoperative organ tracking can be categorized by the method of real-time data collection; ultrasound, x-ray fluoroscopy, and magnetic tracking capture most of the current approaches. King et. al. use 3D ultrasound to perform statistical shape tracking and modeling, while Aylward uses 3D ultrasound and explicit vessel modeling to perform nonrigid tracking. Interventional MR and CT, though still expensive and of limited availability, have also been explored as possible real-time data options by Hartkens et. al [40]. Each of these approaches is discussed in greater detail below.

King et. al. have developed a promising liver tracking method using 3D ultrasound and statistical shape modeling [52]. For each patient, they acquire a series of MR data sets at different points in the respiratory cycle. They next segment the liver and model the boundary

in each data set; this set of boundaries is used to establish a mean shape and eigenmodes of variation for that shape. Polynomial functions are fit to the coefficients of the eigenmodes of variation over the stages of the respiratory cycle, so the final patient-specific model is dependent solely upon the estimated phase within the respiratory cycle. This a priori model is used to constrain a Bayesian registration to the intraoperative ultrasound image data. This approach yielded RMS distance-to-surface errors of about five millimeters on several patient data sets.

King reports that the incremental tracking update takes approximately five minutes on a Sun Ultra 10, so the method is not presently suitable for real-time operation. Nonetheless, improvements in computational efficiency and processing power may bring this promising method within the range of real-time operation in the near future. This method could also be modified to provide a prior model for the balloon-tracking approach described below.

Stephen Aylward has developed a sophisticated set of vessel modeling and registration techniques, including a method for registering preoperative vessel models to the vessels visible in intraoperative 3D ultrasound. Aylward's techniques extract vessels by traversing the intensity height ridges in a 3D image set. Vessel scale is determined using a sub-voxel tubule detection filter, as described in [1–3] . He calculates a coarse rigid registration between the preoperative vessel model and intraoperative ultrasound using mutual information; registration is refined using nonrigid warping of the corresponding vessel tree elements. This approach is computationally intensive, but currently runs at about one update every one to five seconds on a high-end PC. Once the computational challenges are reduced or overcome by improved computing power, this approach holds substantial promise for tracking many types of anatomical structures.

Hartkens et. al. described a method for measuring brain deformation during neurosurgery that could be adapted for following other organs [40]. They perform a non-rigid registration on a field of control points that maximizes the normalized mutual information between two image sets. This method would be similarly applicable to tracking the liver using interventional CT or MR imaging.

Nikos Paragios and Rachid Deriche developed a hybrid of active contours and level set methods that handles complex scenes and changing object topologies ([71] and Paragios'

dissertation [70]). Pixels are segmented into populations of moving and static pixels using mixture modeling, and then an evolving surface is created using the boundaries between those populations as the zero level set of that surface. Calculations on the new surface can then be used to fluidly define changes in the topology of the level set curve families, thus enabling object evolution to occur in a straightforward manner. Their implementation was able to identify the boundaries of four objects in a 200x200 test image in approximately three seconds on a Sun ULTRA 10 at 299MHz. If this method could be accelerated sufficiently for our real-time constraints and altered to enforce a near-elliptical balloon model, it could prove highly effective as an alternative tracking tool in this application.

Corral, Cleary et al have been evaluating the use of a magnetic tracking device (AURORA™, Northern Digital) for intraoperative tracking. In one external test of needle positioning over a 10x10x10cm working volume, they found a mean positional error of 1.22mm [21]. This group has also tested electromagnetic intravascular guidewire tracking as a means of augmenting intraoperative imaging with tracked instrument data [107].

5.1.3 Key Principles of Image-Based Tracking Methods

In order to identify the most promising approaches to a particular tracking problem, it is important to consider some general categories of tracking methods and select those that best accommodate the problem at hand. Two possible categorizations of tracking methods will be considered in this work: local versus regional scale and edge versus body operators. While these categories are admittedly intertwined, they do shed some light on the strengths and weaknesses of each approach.

Local versus regional operators

Some operators consider only local features, such as the intensity difference between adjacent pixels, while others use information from a larger image region, such as an intensity-weighted center of mass (centroid). This concept encompasses a *spectrum* of operators that vary, in general, by the size of their regions of support. Local operations, such as Canny edge detectors, simple thresholding, and active contours with weak prior constraints, tend to be responsive to small local features but highly susceptible to random noise. They also tend to be com-

putationally efficient. On the other end of the spectrum, regional operations like centroids, template matching, multi-element structuring kernels and mutual information measures tend to be more robust in the presence of noise but somewhat more computationally intensive.

Operations in scale space exemplify this spectrum because they can vary from small-scale local features to broad regional features depending upon the scale being considered. This work uses a regional operator with a scale approximately equal to the size of the expected image feature. This approach improves the specificity of the operator for the desired elliptical image feature.

Edge versus body

By definition, edges are contrast-based features that may serve as the basis for tracking an object. However, edges are generally restricted to a relatively small area and are thus subject to degradation by random noise, overlapping objects, and other obscurations of the object's boundary. Edge-based measures may be bound together or applied in scale space in order to form a regional basis for object identification and tracking (as in medial atoms), but such approaches are still dependent upon the relatively thin basis of the object boundary. Object bodies offer a generally larger area over which the object's signal can be differentiated from other noise, but delimiting an object's body from background is a nontrivial task. Choosing between these options requires knowledge of the specific noise characteristics, contrast conditions, and object properties of a given problem.

In general, edge-based measures perform well when object boundaries show good contrast and when relatively few boundary disruptions are expected. Body measures excel when boundaries are unclear or disrupted and overall object position is more important than precise edge localization. This work demands knowledge of object position and orientation and offers little edge contrast or reliability, so a body-based centroid measure was chosen.

5.2 The Intravascular Balloon Catheter as Fiducial

As mentioned above, it is often the case that a structure of interest, such as a vessel or tumor, is difficult or impossible to visualize during the critical phase of a medical procedure. In this

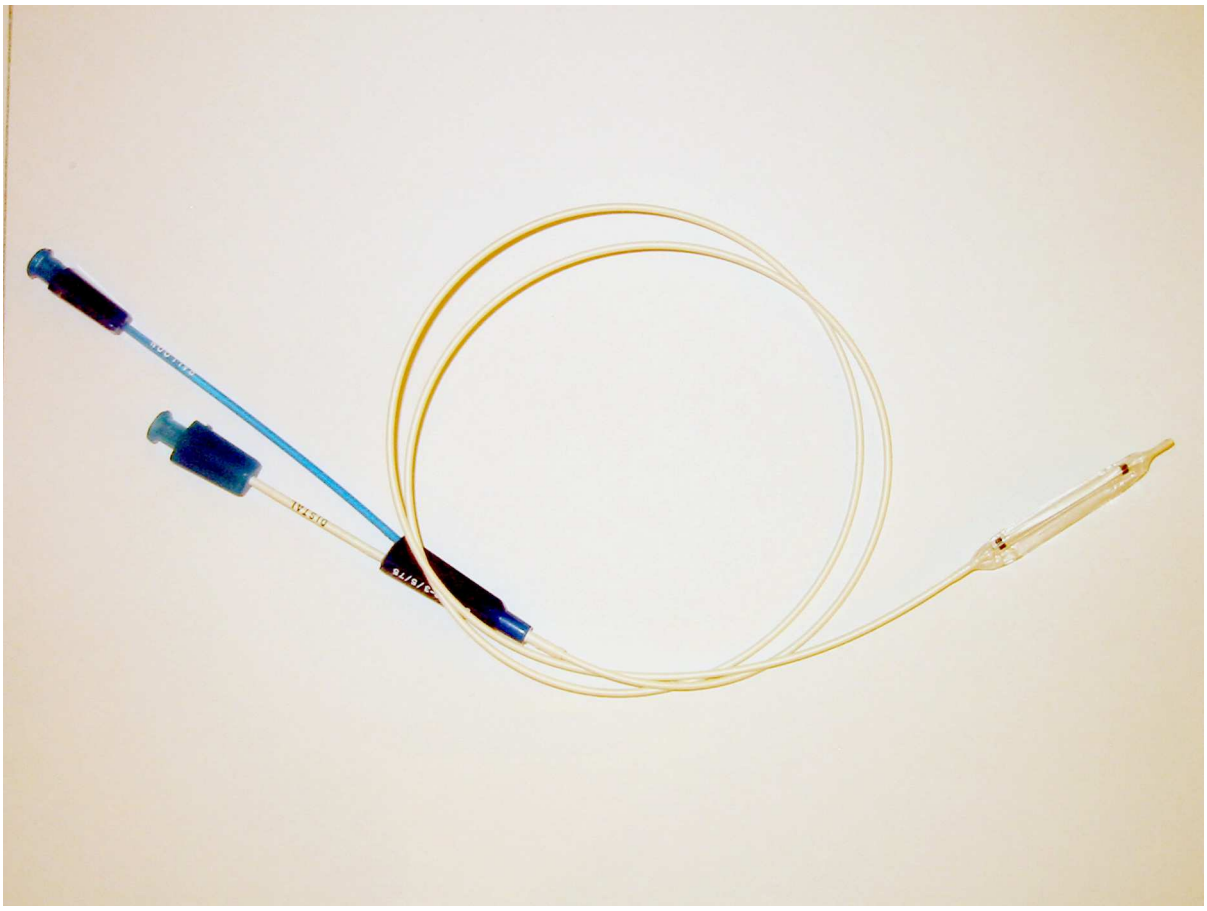
circumstance, one could track a rigidly related target, or proxy fiducial, to obtain a reasonable estimate of the motion of the structure of interest. In this case, a standard intravascular balloon catheter is inflated with a contrast agent and used as such a proxy fiducial[2]. Figure 5.3 demonstrates the physical structure of a standard balloon catheter, while Figure 5.4 shows an array of intraoperative fluoroscopic images of the inflated balloon catheters.

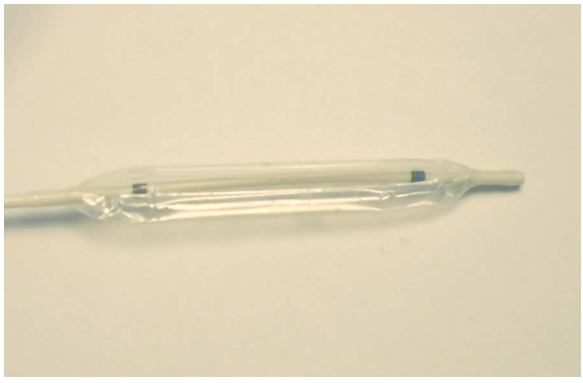
The intravascular balloon catheter is commercially available and commonly used during intravascular procedures. Balloon catheters are available to temporarily slow or block blood flow through a vessel, limit the diffusion of a contrast agent to a branch of a vessel, or, under higher pressure, to forcibly expand a vessel or related structure. In the TIPS procedure, physicians may place a balloon catheter in part of the hepatic vein in order to enhance the diffusion of CO_2 into the portal vein and thus enhance the contrast in the resulting venogram (see §2.1 for details). The common clinical use and FDA approval of these devices has led our clinical colleagues, as well our institution's Committee for the Protection of Human Research Subjects (a.k.a. Institutional Review Board or IRB), to believe that this approach is sufficiently safe for experimental testing in human patients.

In this modified version of the TIPS procedure, *two* access points are created in the right external jugular vein in the neck. An intravascular balloon catheter is placed in one of the right branches of the hepatic vein via the first access point and inflated using an iodinated x-ray contrast agent. The operating physician then performs a portal venogram by injecting dissolved carbon dioxide behind the inflated balloon catheter in order to visualize the nearby portal vein. The physician then introduces the TIPS needle via the second access point, thus allowing both the needle and the intravascular balloon catheter to remain in the liver simultaneously. This means that the balloon can be used to track liver motion throughout the procedure, including the critical phase of advancing the needle from the hepatic vein and into the portal vein. This modified procedure has been performed at UNC under an IRB-approved research protocol with explicit patient consent.

Figure 5.2 shows the typical configuration of the hepatic and portal veins. One may also wish to reference [68, p 273] and [23]; their anatomic descriptions of standard hepatic anatomy

[2] The idea of using a balloon catheter as a tracking fiducial was a joint development among Dr. Stephen Aylward, Dr. Susan Weeks, and the author





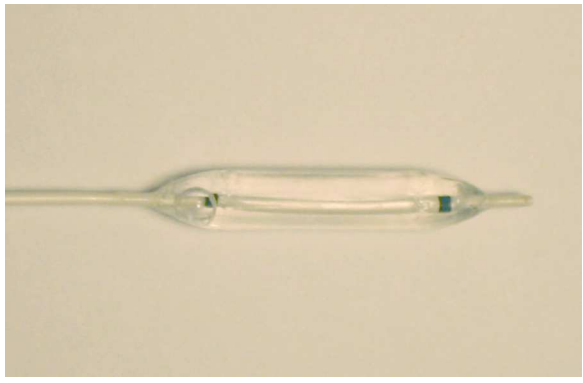


Figure 5.3: Optical images of an intravascular balloon catheter. These catheters are commonly used for non-tracking purposes during interventional procedures.

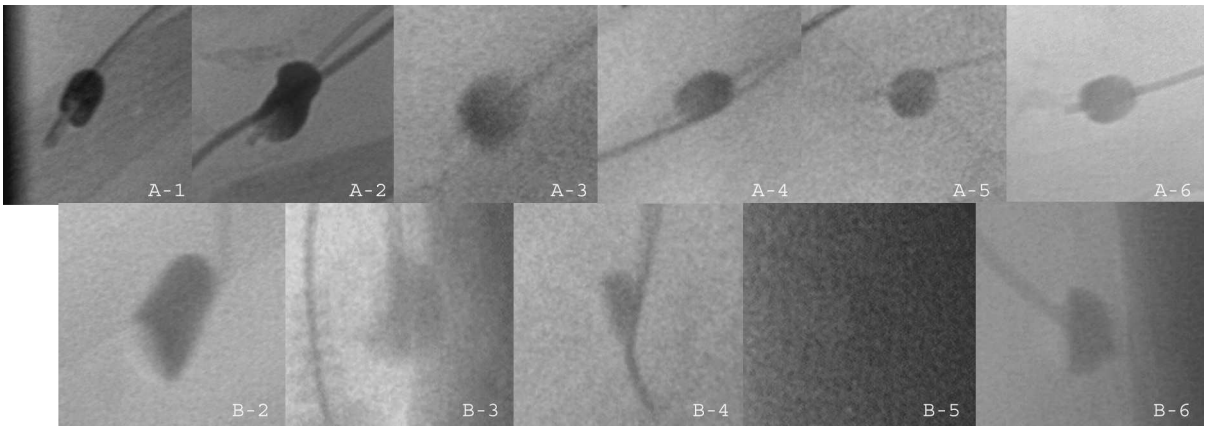


Figure 5.4: Examples of intravascular balloon catheters as seen using x-ray fluoroscopy. The top row of images contains views along the PA (A-x) axis, while the bottom row of images includes views along the lateral (LAT, B-x) axis. Series 1 was only captured from the PA perspective; all others are paired images. All images were gathered intraoperatively under an IRB-approved protocol.

contributed to the design of the aforementioned figure. The intravascular balloon is generally placed in the right hepatic vein approximately 4-5cm from the desired portal vein target.

As shown in Figures 5.3 and 5.4, the balloon catheter takes on a generally ellipsoidal shape when inflated. The balloons have one long axis and two approximately equal short axes. This means that the orientation of the long axis can be determined from two distinct generic views (neither looking straight down the long axis), but it also means that the rotation of the short axes around the long axis is ambiguous. Ellipsoids project to ellipses in all but a few degenerate views[3], so the relative orientation of the long axis of the balloon in the projective plane can be measured as the relative orientation of the elliptical projection. Orientations from two such distinct views provide an estimate of the relative orientation of the balloon (see §5.3.3 for further details). The 3D position of the balloon can be reconstructed from the two

[3] See any discussion of Dandelin spheres for a demonstration of this principle. This approach would fail if the balloon was viewed down the length of its long axis. Fortunately, the anatomical location and orientation of the hepatic vein relative to the cameras make this exceeding unlikely. Any operative system should, however, protect against sudden instability in this metric should a degenerate view occur.

2D centroids of the elliptical projections.

5.3 Computational Methods

The intravascular balloon catheter mentioned above is only one of many features that are visible in intraoperative fluoroscopic images, so it is necessary to differentiate the balloon feature from other distractors before localizing and tracking it. Figure 5.5 shows a small example of the competing image features that are commonly found in the intraoperative views.

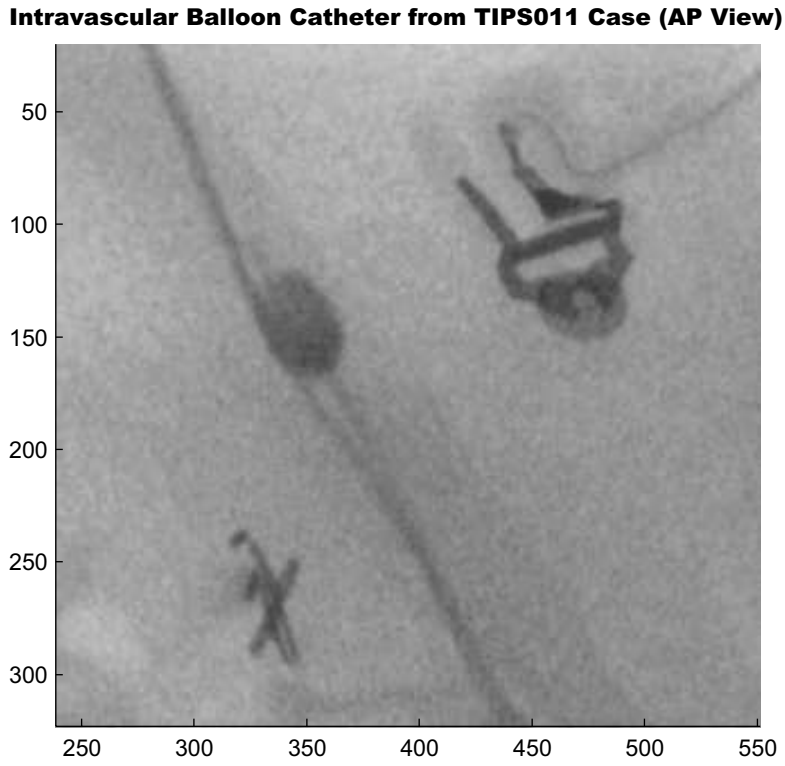


Figure 5.5: Example of an intravascular balloon with possible distractor objects. An electrocardiogram lead and another external metal clip are visible nearby.

The proposed approach to identifying and tracking the intravascular balloon consists of five major components, as follows:

- Feature enhancement (Laplacian of Gaussian filtering)
- Candidate identification and analysis

- Selection of the best candidate (maximum likelihood estimate)
- Reconstruction of the balloon into three dimensions
- Tracking in subsequent frames

In the first step, feature enhancement, the image is filtered with a Laplacian of Gaussian kernel in order to enhance round or elliptical valleys. Next, viable features are identified and analyzed for critical characteristics such as area, eccentricity, and solidity. The best available candidate is then selected as the feature with the maximum likelihood of being the balloon given its critical characteristics. This process is repeated for both views, and then the balloon is reconstructed into three dimensions. Finally, the balloon is followed and the position and orientation of the balloon are updated in subsequent frames from the biplane fluoroscopy system. The update rate ranges from three to thirty Hertz and is set by the operating physician.

5.3.1 Feature Enhancement

An ellipsoidal figure such as our balloon catheter will project onto an ellipse[4], so the first step of this balloon identification method involves a search for such elliptical features. Figure 5.6 shows one such balloon projection with image intensity mapped to figure 'height'. The projection of the balloon forms a distinctive concave, bowl-shaped valley. To take advantage of the predictable shape of balloon's projection, particularly the curvature of the intensity projection, I chose to use a Laplacian of Gaussian filter.

Figure 5.7 shows the shape of the Laplacian of Gaussian kernel. This type of kernel identifies intensity valleys (or hills, depending upon the sign) of similar width while filtering out other types of features. The base Gaussian kernel is defined mathematically as:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{(x-\mu_x)^2}{2\sigma_x^2}+\frac{(y-\mu_y)^2}{2\sigma_y^2}\right)} \tag{5.1}$$

or, in the case of a 2D isotropic Gaussian kernel,

$$g(r) = \frac{1}{2\pi\sigma_r^2} e^{-\frac{r^2}{2\sigma_r^2}} \tag{5.2}$$

[4] unless viewed along the long axis, which is constrained from happening by the imaging geometry and anatomy

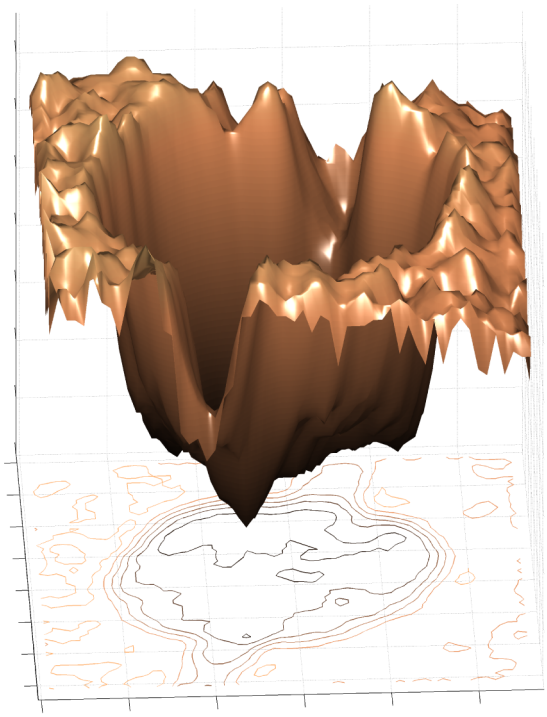


Figure 5.6: Height-mapped view of balloon image. The height of each point in the figure corresponds to the grayscale intensity of this balloon's image at that point. The deep valley represents the balloon, while the two small depressions at the ends represent pieces of the attached catheter.

The Laplacian operator is defined as:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{5.3}$$

The Laplacian of Gaussian filter is thus described as:

$$LoG(x,y) = \nabla^2 g(x,y) \tag{5.4}$$

$$= \frac{\partial^2 g(x,y)}{\partial x^2} + \frac{\partial^2 g(x,y)}{\partial y^2} \tag{5.5}$$

$$= -\frac{1}{2\pi\sigma_x\sigma_y}\left(\frac{1}{\sigma_x^2} + \frac{1}{\sigma_y^2} - \frac{(x-\mu_x)^2}{\sigma_x^4} - \frac{(y-\mu_y)^2}{\sigma_y^4}\right) e^{-\left(\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2}\right)} \tag{5.6}$$

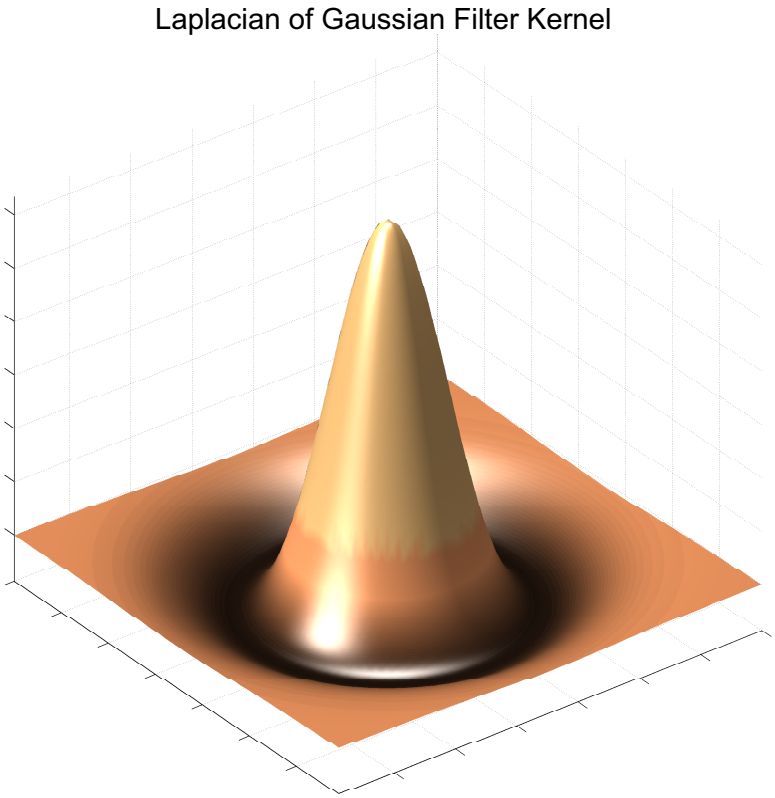


Figure 5.7: Two-dimensional Laplacian of Gaussian filter kernel used in the balloon detection algorithm. The filter is inverted (upside-down) so that the filtered balloon features remain darker than the background.

or, in the case of an isotropic Gaussian,

$$LoG(x,y) = -\left(\frac{1}{\pi\sigma^4} - \frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\pi\sigma^6}\right) e^{-\left(\frac{(x-\mu_x)^2+(y-\mu_y)^2}{2\sigma^2}\right)} \tag{5.7}$$

This can be equivalently described as a function of radial distance r,

$$LoG(r) = -\left(\frac{1}{\pi\sigma^4} - \frac{r^2}{2\pi\sigma^6}\right) e^{-\left(\frac{r^2}{2\sigma^2}\right)} \tag{5.8}$$

which has its finite zero-crossings at

$$\begin{aligned} 0 =& \frac{1}{\pi\sigma^4} - \frac{r^2}{2\pi\sigma^6} \\ r =& \sigma\sqrt{2} \end{aligned} \tag{5.9}$$

When the edges of an object are aligned with the zero-crossing boundary of the filter, then the filter effectively measures the contrast difference between the data in its central valley and the surrounding background. When this is the case, one can argue that the filter is at the correct size or scale for that object. If the scale of the Laplacian of Gaussian filter is selected to match the expected width of the balloon catheter's image, then the filter will respond optimally to a dark balloon of that size embedded in a brighter background. Equation 5.9 shows that choosing scale $\sigma = \frac{1}{2\sqrt{2}}w$, where w is the expected width of the balloon's image, should match the filter correctly to the size of the balloon.

There are, however, concerns other than maximum filter response in the selection of the filter's scale. The true goal of the filter is to maximize the signal difference between the balloon feature and all non-balloon features. Beneficially, the filter not only enhances the signal from objects near the selected scale, but it also *suppresses* objects of other scales or of inverted contrast. In this particular problem, this effect is most important in the suppression of the catheters and EKG leads that are often located near the intravascular balloon. These objects are generally narrower (of smaller scale) than the balloon, so it may be desirable to select a filter of slightly larger scale than suggested above in order to increase the suppression of the small scale objects. Further experimentation will be necessary to determine the best tradeoff between balloon enhancement and non-balloon suppression.

At initialization, the Laplacian of Gaussian filter is convolved with each of the fluoroscopic images. The effects of this filter are shown in Figure 5.8. The following section describes how these filtered images are used to identify candidate features and select the most likely balloon feature.

Identifying Candidate Features

Candidate features are identified via a series of image analysis and processing steps. First, the image is filtered using the Laplacian of Gaussian filter mentioned above. The image is then thresholded, features are labelled using connected components, and the area, centroid, orientation, perimeter, solidity, and eccentricity of each feature is calculated. These metrics will be used to identify the most balloon-like feature from among the set of all possible features. The following sections describe some of the non-trivial calculations for these metrics.

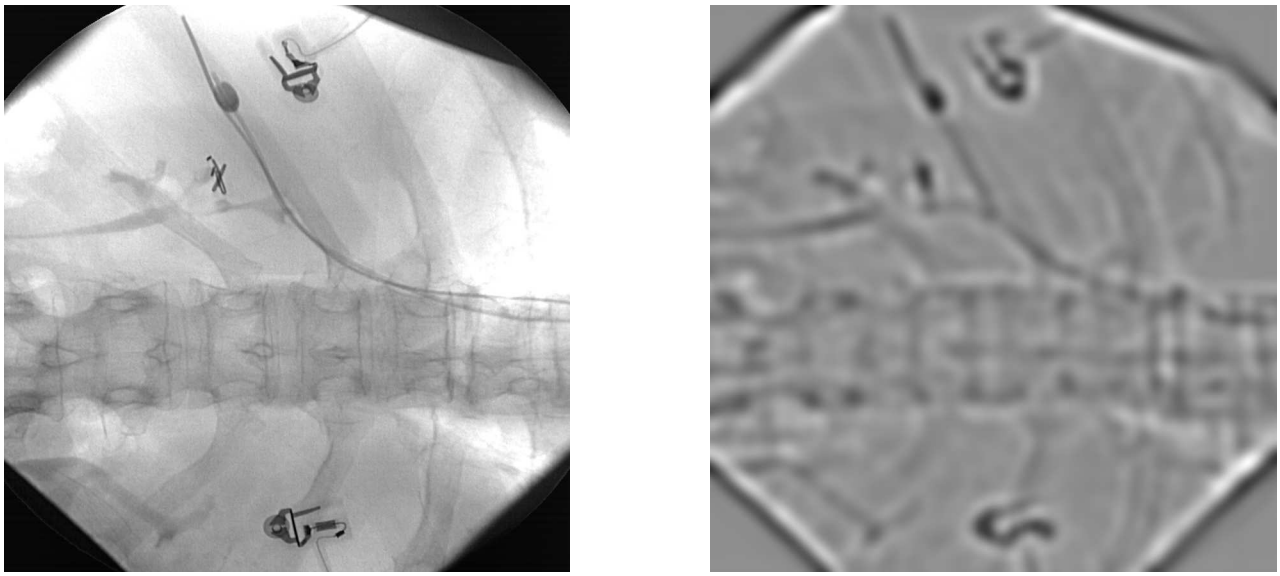


Figure 5.8: Effects of Laplacian of Gaussian Image Filter. The left image shows the unfiltered image, while the right image shows the results of Laplacian of Gaussian filtering. The elliptical balloon image is enhanced while the catheter, EKG clips and other undesirable features are suppressed.

Centroid & Moments

Before one can use the system's calibrated projections to perform 3D reconstruction, it is necessary to be able to accurately calculate the coordinates of the target of interest from noisy images. One robust approach to determining the coordinates of a target is to calculate the center of intensity mass, or centroid, of the target in the image. This is the approach used in Bar-Shalom's paper on the subject, which will serve as the basis for the centroid discussion below [7].

First Moment — Centroid

The location of the center of intensity mass, or centroid, of an object is defined as:

$$\vec{\omega}_c = \frac{\sum_{i,j} \vec{\omega}_{ij} I_{ij}}{\sum_{i,j} I_{ij}} \tag{5.10}$$

where $\vec{\omega}_c$ is the calculated coordinate vector of the centroid, $\vec{\omega}_{ij}$ is the coordinate vector for the pixel at (i, j), and I_{ij} is the image intensity of the pixel at (i, j). The pixel intensity can

be more precisely specified by:

$$I_{ij} = s_{ij} + \phi_{ij}$$

where s_{ij} is the pure signal intensity and ϕ_{ij} is a random noise term, both measured at pixel (i, j). The noise term is assumed to be zero-mean and white with standard deviation σ_p. Bar-Shalom provides a detailed analysis of the contribution of image noise to the centroid calculation above. He shows that:

$$\hat{\vec{\omega}}_c = \vec{\omega}_c + \vec{\omega}_c \, \epsilon_1 + \epsilon_2 + \epsilon_3 \tag{5.11}$$

with $\hat{\vec{\omega}}_c$ defined to be the true centroid position, ϵ_1 to be the multiplicative noise in the measurement, ϵ_2 as the additive noise in the measurement, and ϵ_3 as the error due to image discretization. These three error terms are all assumed to be zero-mean and white (uncorrelated) with Bar-Shalom's derived variances as follow:

$$\sigma_1^2 = \frac{m\sigma_p^2}{s^2 + m\sigma_p^2} \tag{5.12}$$

$$\sigma_2^2 = \frac{\sigma_p^2 \sum_{i,j} \vec{\omega}_{ij}}{s^2 + m\sigma_p^2} \tag{5.13}$$

$$\sigma_3^2 = \frac{1}{12s^2} \sum_{x,y \in T} s_{xy} \tag{5.14}$$

$$= \frac{1}{12m_T^2} \quad \text{assuming constant } s_{xy} \tag{5.15}$$

where s is the sum of the target-related intensities, m is the number of pixels in the image, and m_T is the number of pixels covered by the target. T is the set of pixels belonging to the target. This error model yields a state-dependent term, which will become important in the development of the Kalman filter for this task.

Second Moment - Optimal Ellipse

It is also necessary to determine the orientation of the balloon catheter feature for our catheter-tracking task. Due to the approximately elliptical shape of the balloon catheter on projection, we can use the second central moment of our target's intensity data to calculate the shape and orientation of our target's best-fitting ellipse. The algorithm for making these calculations is

described in [38]. The outputs of this algorithm are the major and minor axis lengths of the best-fitting ellipse and the angle between the major elliptical axis and the horizontal axis.

5.3.2 Selection of the Best Candidate Feature

Consider Ψ to be the set of candidate features in a given image. Our task is to identify the feature $\psi \in \Psi$ that is most likely to be the true balloon feature. This is analogous to the traditional MLE problem of calculating a mean, variance, or other parameter that maximizes the likelihoood of the given samples; in this case, the choice of the true balloon feature serves as the parameter of interest and the various quality metrics serve as the samples.

The first step in the MLE process is to define likelihood distributions for each of the observed metrics. For each metric Ω we define a distribution $l(x)$ where x is an observed value of Ω and $l(B \mid \omega)$ is the likelihood that the feature is a balloon catheter given the observed metric value ω. No empirical distributions were available for this task, nor was the number of available images sufficient to reasonably characterize the true distributions of these metrics across true features and distractors. In lieu of empirical metrics, distributions were chosen based upon the observed parameters of a set of observed balloon and non-balloon features from the limited data sets available. To compensate for the likely exclusion of reasonable metric values using this approach, the selected distributions were assigned a minimum likelihood of 0.05 (5%) across all values of the metric. The maximum value of the base distribution was normalized to 0.60 (60%) prior to the inclusion of the robust floor. Such 'robust' distributions are nicely reviewed by Triggs et al [99]. In the future, it would be better to develop empirical distributions using large and diverse balloon-tracking data sets.

Area

Area is useful metric for excluding features that are much too small or too large to be a balloon catheter. Due to the differences in the apparent zoom between the PA and lateral views, the area distributions for the two views were defined separately. For the PA view, the likelihood function was defined as the maximum of the robust floor of 0.05, as described above, or the gaussian PDF with mean 1500 $pixels^2$ and standard deviation of 400 $pixels^2$. For the lateral distribution, the mean was selected to be 2000 $pixels^2$ with the same standard

deviation as the PA view. The larger area on the lateral view reflects the narrower field of view (and accompanying zoom effects) associated with that view. These distributions were chosen based on the observed values in the few available image series; they are not based on actual measured distributions since insufficient data were available. Due to variations in the size of the balloons used in the OR and different zoom settings on the fluoroscopy system, it was necessary to adjust the mean values of this metric to accomodate actual variations in the areas of the balloons in different series.

Solidity

Solidity is defined here as the fraction of pixels within an object's convex hull that are also within the object. Perfectly convex objects have 100% solidity, so an elliptical balloon, which is convex by definition, should also have a 100% solidity. The gaussian cumulative distribution function (cdf) was chosen as the basis for this distribution because it is monotonically increasing and would reach a maximum at 100% solidity. This work used a gaussian cdf with mean 0.95 and standard deviation 0.05. These distributions were chosen based on the observed values in the few available image series; they are not based on actual measured distributions since insufficient data were available.

Eccentricity

The eccentricity of a feature is defined as the ratio of lengths of the minor to major axes of the best-fitting ellipse. Eccentricity is a measure of how elongated an object is; lines have an eccentricity of zero, whereas circles have an eccentricity of one. Ellipsoidal balloons should be somewhat closer to circles than they are to lines, but should not lie completely at either extreme. The distribution for eccentricity was based upon a gaussian pdf with mean 0.8 and standard deviation 0.6, which corresponds to an ellipse that is 20% longer in its major axis than in its minor axis. These distributions were chosen based on the observed values in the few available image series; they are not based on actual measured distributions since insufficient data were available.

Area:Perimeter ratio test

The ratio of an object's area to its perimeter describes the 'roundness' or area efficiency of that object[5]. A tortuous or linear object will have a low area relative to its perimeter, whereas an elliptical feature will have a high area relative to its perimeter. The values of this ratio range from zero for arbitrary curves to $\frac{r}{2}$ for a circle of radius r.

The expected ratio for a given feature can be calculated based upon the major- and minor-axis lengths of the best-fitting ellipse for that feature. An axis-aligned ellipse is defined by the equation $x^2/a^2 + y^2/b^2 = 1$, where a is the half of the major axis length and b is half of the minor axis length. The area of such an ellipse is given by $A = \pi ab$. The perimeter of the ellipse is closely approximated by $P = \pi\sqrt{2(a^2 + b^2)}$. This means that the ideal area-to-perimeter ratio for an ellipse is $R = \frac{A}{P} = \frac{ab}{\sqrt{2(a^2+b^2)}}$.

The likelihood distribution for this metric is actually based upon the normalized area:perimeter ratio, or the ratio between the measured and expected area:perimeter ratios. Due to the way in which the perimeter is measured in the pixelized images, the perimeter tends to be slightly underestimated, yielded consistently higher area:perimeter ratios than predicted. The likelihood distribution therefore was based on a gaussian cdf with mean 1.1 and standard deviation 0.1 because the distribution is monotonically increasingly (and higher A:P ratios correspond to more circular features) and the maximum slope (at 1.1) occurs before the observed value of the A:P ratio in the ellipses in this series. These distributions were chosen based on the observed values in the few available image series; they are not based on actual measured distributions since insufficient data were available.

Minimum Value

The minimum value of a feature is the maximum negative filter response that is included within that object. This can also be thought of as the greatest 'depth' of the intensity valley for the intensity surface corresponding to that feature. While this value varies from image to image and balloon to balloon, it reflects a sense of how well a particular feature matched

[5] Despite extensive searches, the author was unable to find an original reference for the use of the area-to-perimeter ratio as a metric of roundness. It appears frequently in discussions of feature discrimination without citation.

the scale and shape of the selected Laplacian of Gaussian filter. The likelihood distribution for this metric was based on the function $1 - \mathbf{cdf}(-0.10, 0.06)$, where cdf() is the gaussian cdf function, -0.10 is the mean, and 0.06 is the standard deviation of the distribution. This distribution was chosen because it is monotonically decreasing (deeper, more negative valleys are more likely to be balloons) and the maximum slope occurs between the average balloon value and that of most distractors. These distributions were chosen based on the observed values in the few available image series; they are not based on actual measured distributions since insufficient data were available.

Position

The intravascular balloon catheters are generally placed in similar locations in each case, so some value may be assigned to the proximity of each feature to the typical location of a balloon in that view. The likelihood distributions were distinct for each view. For the PA view, the distribution was based on a 2D gaussian pdf with mean coordinates (500, 200) and standard deviation 75 pixels. For the lateral view, the mean coordinates were (440, 600) with standard deviation 120 pixels. For subsequent iterations, the expected coordinates were set to the prior coordinates of the feature centroid and the standard deviations were reduced to 20 for the two views. These distributions were chosen based on the observed values in the few available image series; they are not based on actual measured distributions since insufficient data were available.

Composition of Likelihoods

Having assigned likelihoods for each of the metric values described above, it is now necessary to combine these likelihoods into an overall likelihood that a given feature is a balloon catheter. For simplicity, I chose to use the product of the likelihoods of the metrics for this purpose. Define M as the set of all metrics Ω, ω_ψ as the value of a given metric Ω for a candidate feature ψ. The overall likelihood that ψ is a balloon is then:

$$L(\psi \text{ is a balloon}) = \prod_{\Omega \in M} l(B|\ \omega_\psi \text{ in } \Omega) \tag{5.16}$$

The best candidate is chosen as the $\psi \in \Psi$ that maximizes $L(\psi \text{ is a balloon})$.

Other choices of combination functions do exist, particularly in the field of fuzzy logic. It may be valuable to explore the benefits of such methods in future iterations of this work.

5.3.3 Balloon Reconstruction into 3D

Once the balloon feature has been identified in both frames, the 3D position of the intravascular balloon catheter is reconstructed using the methods described in section 4.3.2. In addition, the orientation of the balloon is estimated from the camera calibration and from the estimated balloon axes in each view. An iterative equation solver (Matlab's fsolve function) is used to find the orientation around the reconstructed centroid that best coincides with the measured elliptical axis orientations. Due to the ambiguity of orientation around the main axis of the elliptical balloon, only two angles of rotation can be determined. The third is assigned arbitrarily and assumed to remain constant throughout the procedure.

5.3.4 Tracking in Subsequent Frames

After the initial balloon match has been made, it is reasonable to restrict the subsequent searches using our prior knowledge of the balloon position and characteristics. Two techniques are used in this work: restricting the search region to a small area near the prior balloon position, and adjusting the feature likelihood distributions to reach maxima at the previously measured values for the balloon. The first measure improves the processing efficiency of the system dramatically; the second improves the robustness of the system to changes or distracting features in the images.

5.3.5 Initial Results

Figure 5.9 demonstrates the performance of the system in one example series. The tracker was able to distinguish among several possible features successfully. The next chapter presents results over several series, including some with poor image quality or numerous distractor objects.

Variations in Camera Settings, Balloon Size, and Air Spaces

The most significant problem in this tracking work was that the balloon catheter image was not well standardized between cases. Three major sources of variation were encountered:

1. Balloon size - Different balloon catheters were used in some of the cases; some balloons were long, some were short, and some were in between. These differences yielded differences in the area, area:perimeter ratio, and eccentricity metrics.

2. Air spaces - The balloons had air spaces in some cases, which altered the detected area, area:perimeter ratio, and the eccentricity of the ballons.

3. Camera settings - The focal lengths of the cameras were frequently different between cases, which yielded substantial differences in the area and position of the balloon catheter.

In future research, it should be possible to standardize the first two items above across cases. It may be imperative that a standardized balloon size and type be used consistently until sufficient data is available to model additional catheter styles accurately. The third item, camera settings, will need to be managed within the tracking code. Such differences in views are often necessary to manage variations between patients and thus must be managed as part of the medical needs of the procedure.

5.4 Future Work

The most important next step in this work is the development of empirical likelihood distributions for the various metrics. In addition, the relative discrimination power of the metrics should be determined and any underperforming metrics should be eliminated. The development of empirical distributions should also include the consideration of the projection parameters of the two cameras (zoom correction) as well as the size of the balloon catheter being used so that predictable variations in the balloon features between cases can be handled robustly. It should not be necessary to routinely 'tweak' the area distributions to manage routine changes in projective zoom or balloon catheter size.

The position metric showed less than optimal reliability on some of the initial frames but substantial value in subsequent frames. Part of the difficulty with the initial identification task is that the distribution of balloon positions is not modeled well by a gaussian pdf. It is almost certainly necessary to use an empirical distribution for this metric, at least for the initial search for the balloon feature. Subsequent iterations may be well modeled by a gaussian distribution due to the improved prior knowledge of the balloon position and its pseudo-random motion within the local region around its prior position. Operator intervention is a possible option should automated methods fail, but such intervention would reduce the effectiveness of the system by requiring a dedicated operator and limiting the system from automatically recovering from errors.

The original implementation of this tracking system was done using the Matlab development environment, which made implementation relatively straightforward but provided less-than-ideal performance. One of the other students in the laboratory, Nick Carruthers, ported this implementation over to OpenCV and found a 70-fold improvement in speed. Using this implementation, search updates over a 256x256 region take only 23ms. Since the fluoroscopic images are updating at 30Hz or less, this will allow sufficient time for balloon tracking to occur in each frame before the next frame becomes available. Additional benefits may be garnered from optimizing the selection of region-of-interest windows to lower the average window size.

There may also be more robust ways to identify candidate features in the image. While the current filtering and thresholding scheme has yielded good results to date, it does not take full advantage of the geometric properties of the balloon projection. In particular, it does not fully utilize the convexity of the image intensity of the balloon feature. It may be possible to measure special points in image features, such as the rim of parabolic points around a region of convexity, to provide a more robust estimation of the balloon boundaries. This might eliminate the erroneous linkage between similar-intensity regions that can occur when thresholding closely opposed features. The computation intensity, however, is likely to be significant.

It may also be tempting to use the association of the balloon to an intravascular catheter as an additional identifier. While this association would always be present, it is likely that the large number of tubes, wires, and cords on the average patient would present too many

distractions to make this metric reliable. It may be possible to use catheters with distinctive markings (such as measurement catheters with rings every 1cm) to discriminate the true catheter from all other distractors.

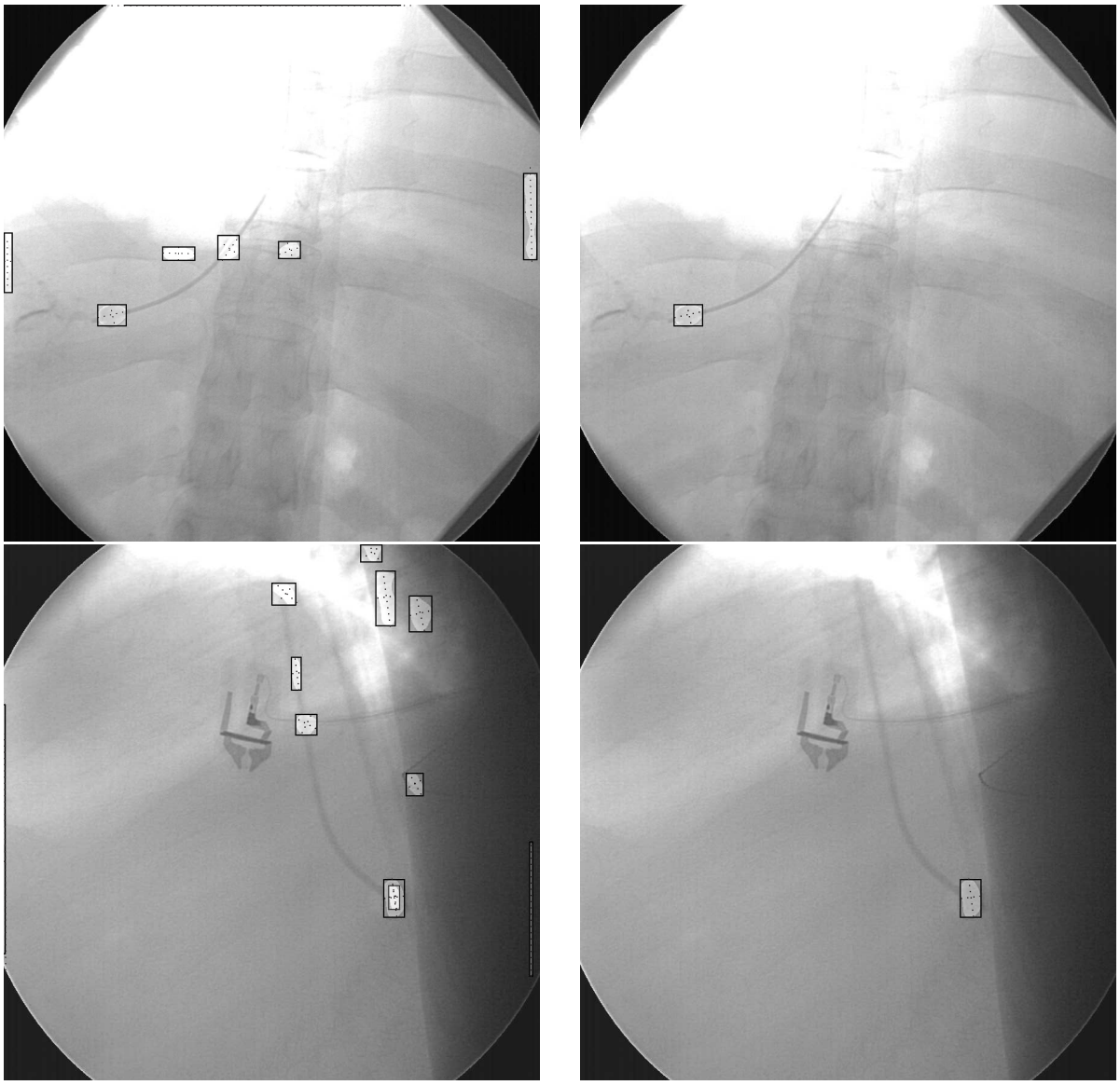


Figure 5.9: Images of the initial results of the balloon selection system. The images on the left show all of the balloon candidates identified in the image, whereas the images on the right show only the selected candidate. The top row shows an PA view; the bottom row shows a lateral view.

Chapter 6

TRACKED LIVER MOTION DURING TIPS

The previous chapter described methods to identify and track a proxy of intraoperative liver motion. In this chapter, these methods will be applied to several sets of intraoperative image data to determine the motion of the liver during the TIPS procedure.

Relatively few studies of liver motion have been published. Levy and colleagues used ultrasound to measure liver vessel motion during respiration and found motion of 11mm during the respiratory cycle [56]. Clifford and colleagues published an excellent review of liver motion studies in 2002 [18]. They found that nine articles on the subject reported motion of 10-28mm in the craniocaudal direction during quiet respiration; most studies reporting motion of about 10-14mm. Motion in other directions is generally reported as less than half the magnitude of any craniocaudal motion. Korin and colleagues studied abdominal organ motion during respiration using MRI and found an average of 13mm of craniocaudal motion over the breathing cycle [53]. Davies et al used ultrasound to measure respiratory motion and found an average of 10 and 37mm of craniocaudal motion during shallow and deep respiration, respectively [22]. Herline et al used optical tracking to measure liver motion for stereotactic surgery; they found an average of 10.8mm of craniocaudal liver motion [43]. Shimizu used MRI to demonstrate mean motion of about 11mm craniocaudally and approximately 5mm in the anteroposterior and lateral directions [88, 89]. Kubo and Hill reported liver motion of 1.5 to 2.0cm during respiration [54]. Venkatraman, Bullitt, and colleagues studied liver motion during respiration and needle pressure in a single patient series using the image acquisition system described in this work[1]. They found 3.9-12.3mm of rostrocaudal motion during respiration and 2.7-13.2mm of rostrocaudal motion during needle pressure [104].

[1] Different calibration and tracking methods were used in this work than in Venkatraman's paper.

6.1 Experimental Design

The data for this experiment were acquired under an IRB-approved research protocol with written patient consent. Four series of images were captured from four separate patients. For each patient, an intravascular balloon catheter was placed in a hepatic vein branch using the methods described in the prior chapter. A short sequence of biplane fluoroscopic images was then captured over one full respiratory cycle. In three of the series, short sequences of images were also captured during attempts to pass the TIPS needle from the hepatic to the portal vein. The images were captured to our PC-based capture system with an anonymous patient identifier. The image series were analyzed offline using the tracking methods described in the prior chapter.

Preliminary results from this work were presented by the author in a prior work [82].

6.2 Liver Motion during Respiration

The liver abuts the diaphragm, so the liver is known to move during inspiration and expiration. By tracking the intravascular balloon catheter over a respiratory cycle, we can estimate the extent of motion throughout these cycles. The following section describes the results of such motion analysis for five respiratory image series captured in four patients. The TIPS 010, 012, and 013 series each include a single series of respiratory motion images. The TIPS 011 series includes one short series and one longer respiratory series. The TIPS012 series was excluded from the lateral results due to inadequate visualization of the lateral balloon (obscured by tissue of arms and spine).

Figure 6.1 includes graphs of the position and orientation variables over each patient series, while Figure 6.5 shows histograms of the displacements across all series. Figure 6.7 shows a visual trace of balloon paths over a respiratory cycle. The tracking results from these series suggest that most of the respiratory motion occurs in the craniocaudal direction (6mm range), with relatively less motion occuring in the lateral and anteroposterior directions (2mm and 3mm ranges, respectively). The per-frame tracking results correlated well with the visual position and orientation of the balloon in each frame and in both views. These results are consistent with those of previously published studies [18].

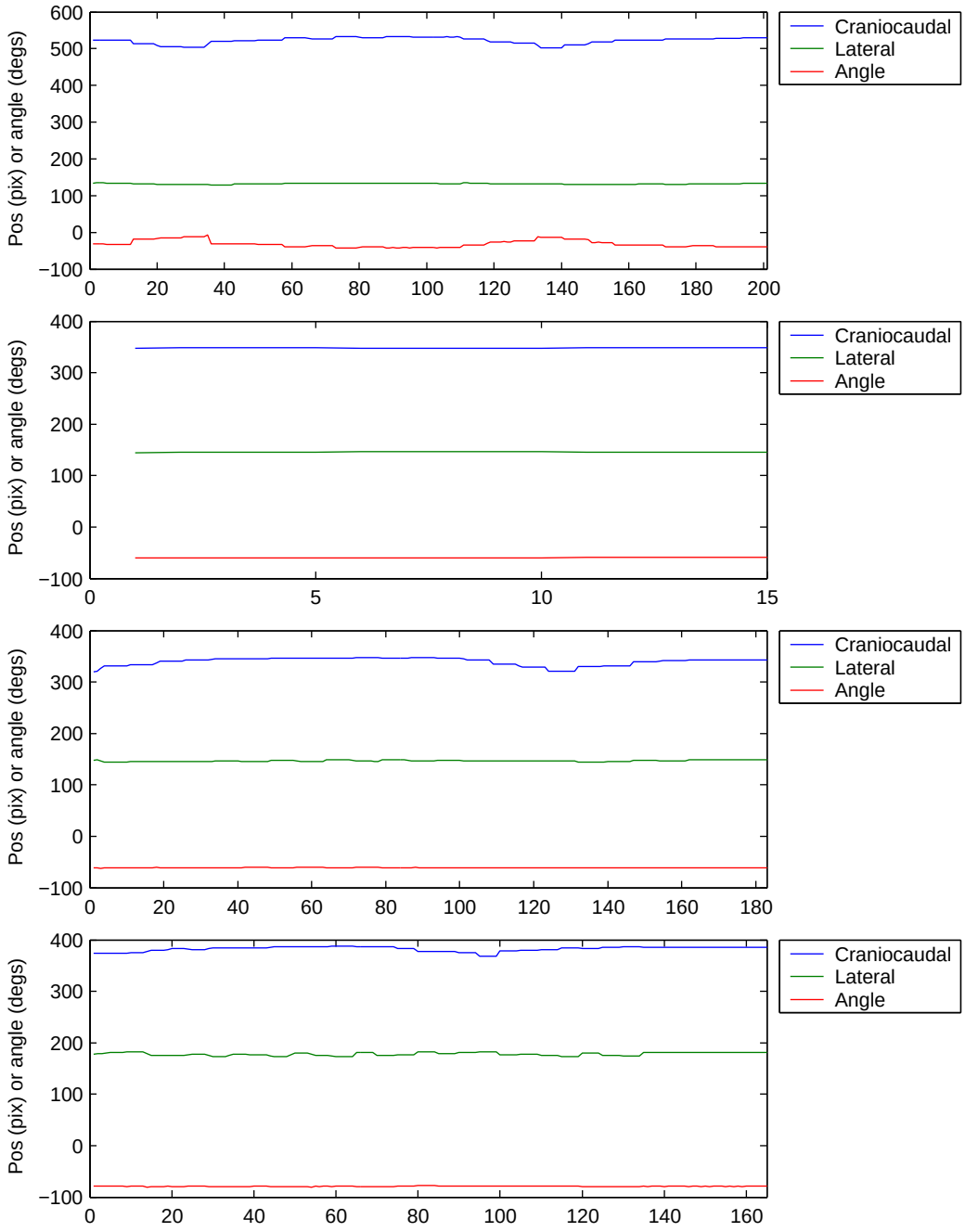


Figure 6.1: Motion of the intravascular balloon on PA images during respiration. The series shown from top to bottom are TIPS010, TIPS011-1, TIPS011-2, and TIPS013.

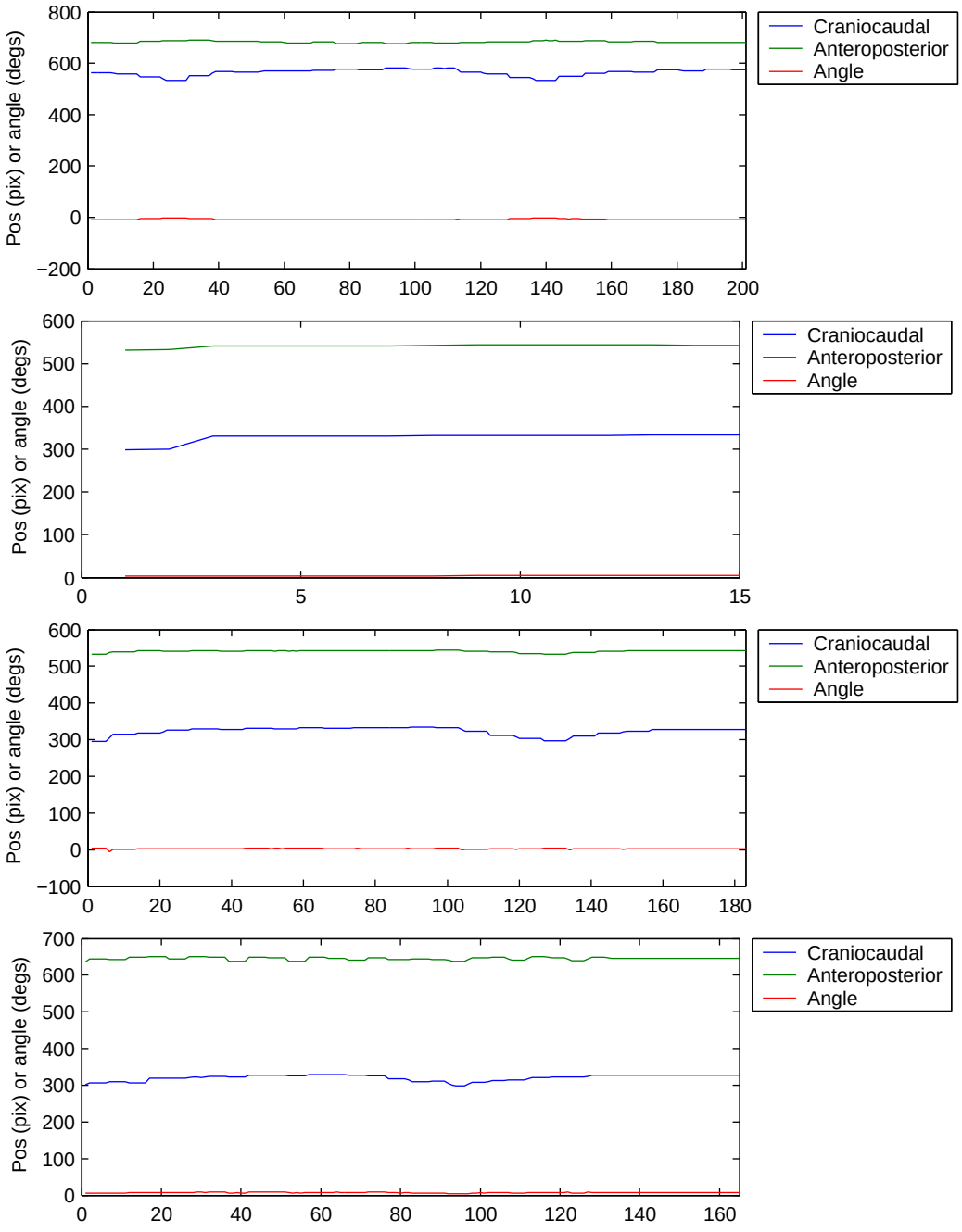


Figure 6.2: Motion of the intravascular balloon on lateral images during respiration. The series shown from top to bottom are TIPS010, TIPS011-1, TIPS011-2, and TIPS013.

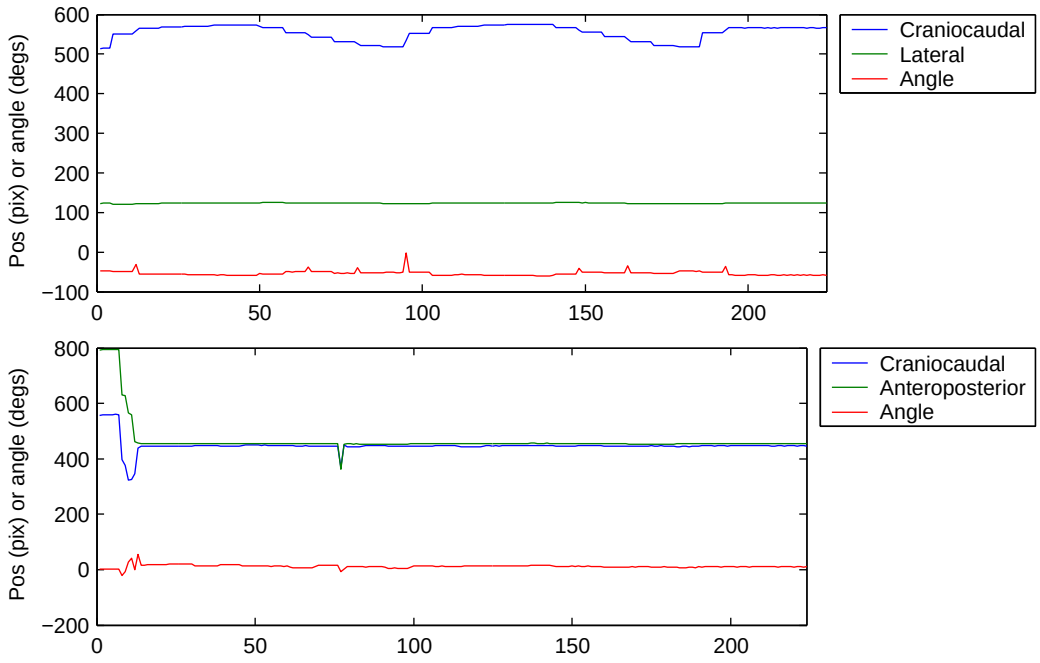


Figure 6.3: Tracked motion of the intravascular balloon during failed TIPS012 series. The PA (upper) chart shows the successful track of the visible view of the balloon, while the lateral (lower) chart shows the tracker failing as the balloon disappears from view. The balloon was not visible to the human eye on the lateral frames.

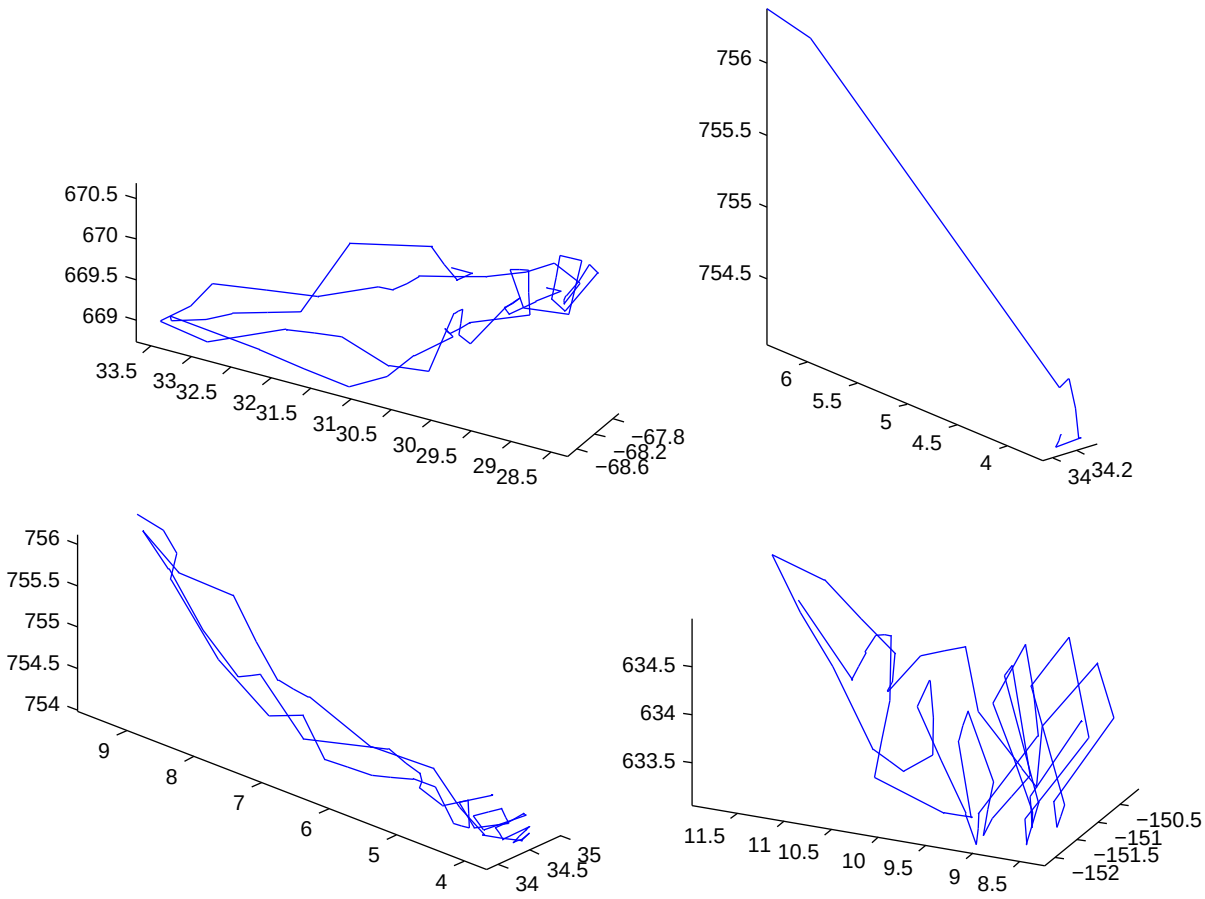


Figure 6.4: 3D motion of the intravascular balloon during respiration. The series shown from top to bottom are TIPS010, TIPS011-1, TIPS011-2 and TIPS013. Units are in millimeters.

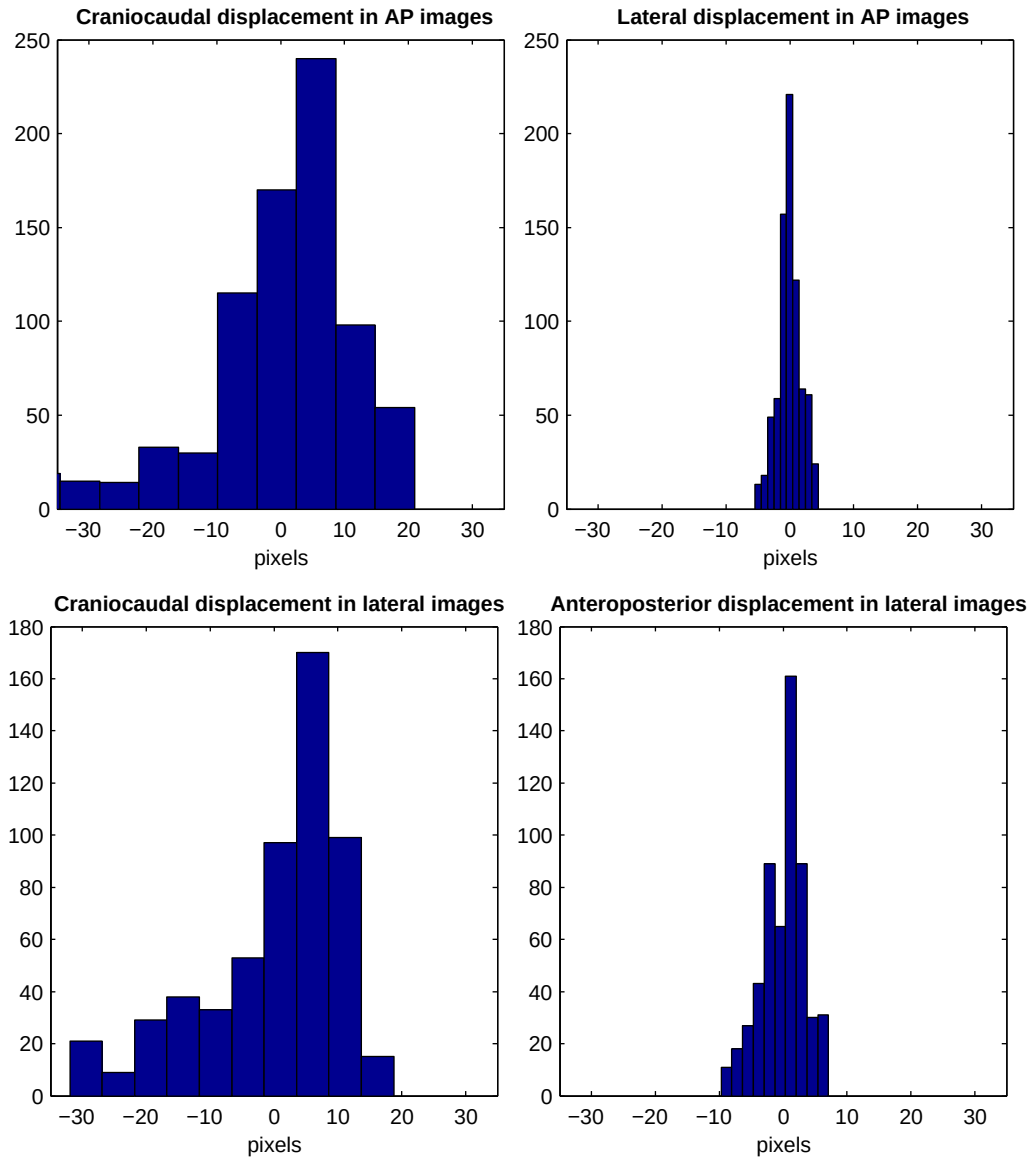


Figure 6.5: Distributions of displacements during respiration. The top row shows displacements in the PA images in the craniocaudal (left) and lateral (right) directions. The bottom row shows displacements in the lateral images in the craniocaudal (left) and PA (right) directions. Each histogram represents the aggregate results from all of the series (TIPS010-TIPS013). The TIPS012 series was excluded from the lateral results due to inadequate visibility of the lateral balloon.

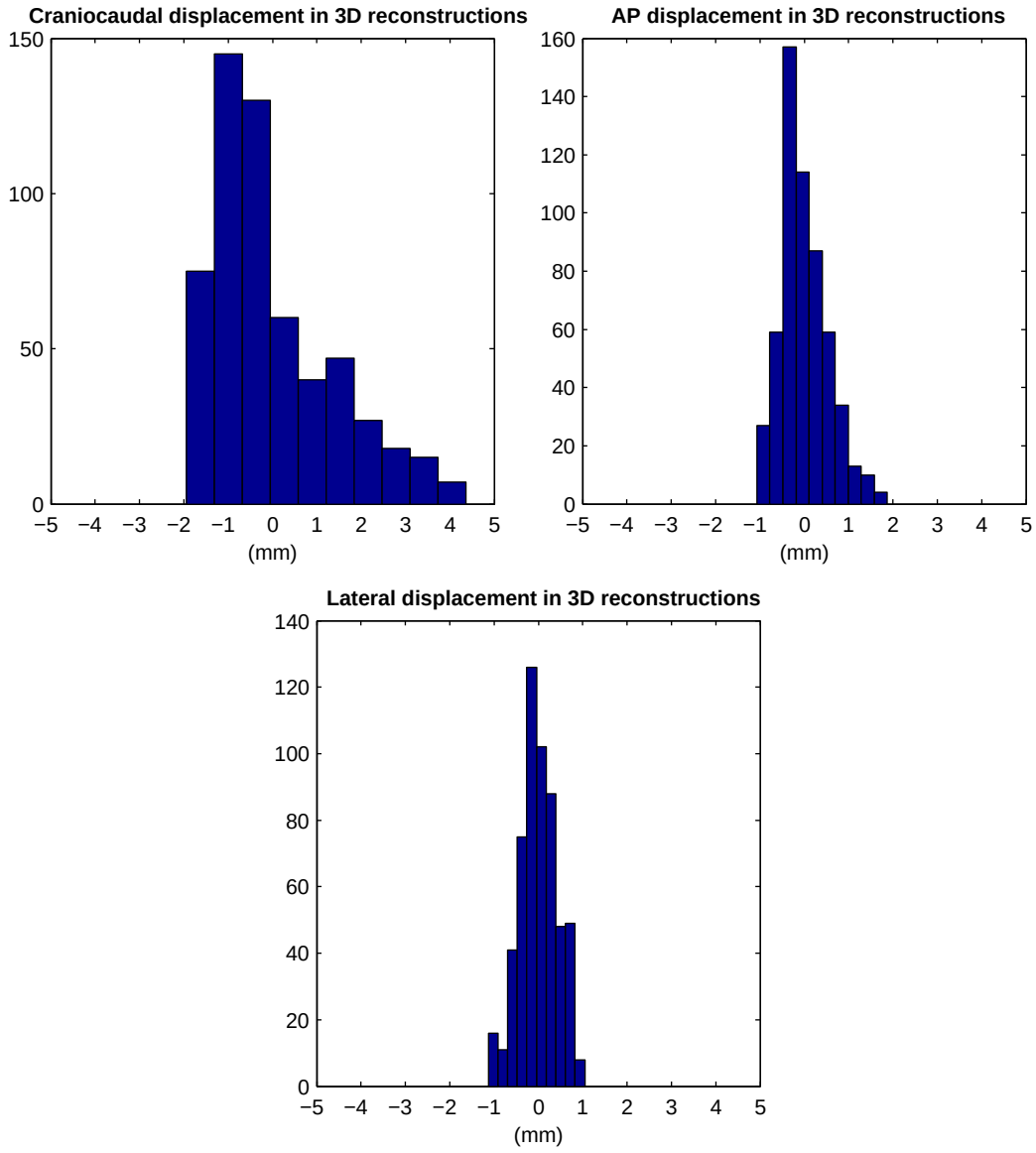


Figure 6.6: Distributions of 3D displacements during respiration. The graphs show 3D displacements in the craniocaudal (top left), PA (top right) and lateral (bottom right) directions. Each histogram represents the aggregate reconstruction results from all of the viable series (TIPS010, TIPS011, and TIPS013).

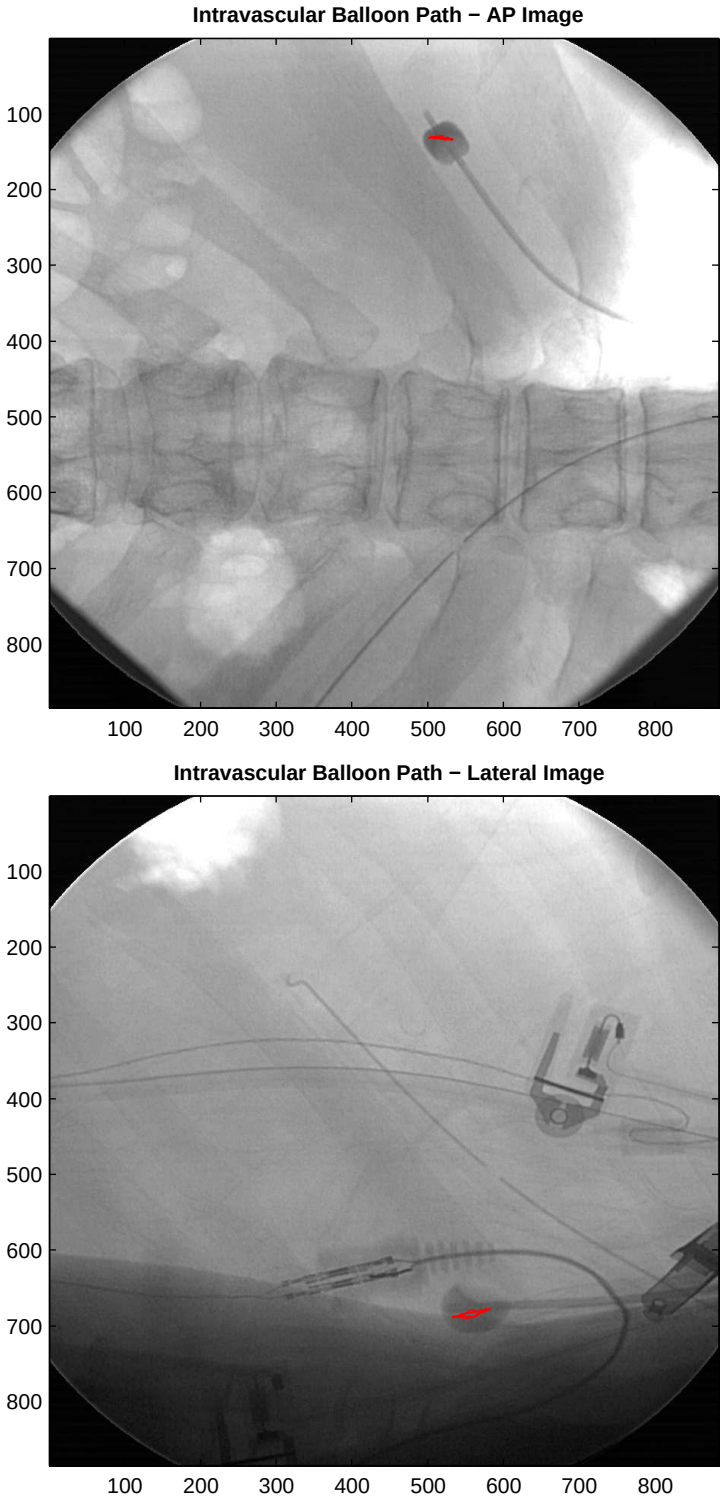


Figure 6.7: Balloon motion during TIPS010 breathing series. These images show the balloon motion during a respiratory cycle (PA top; LAT bottom).

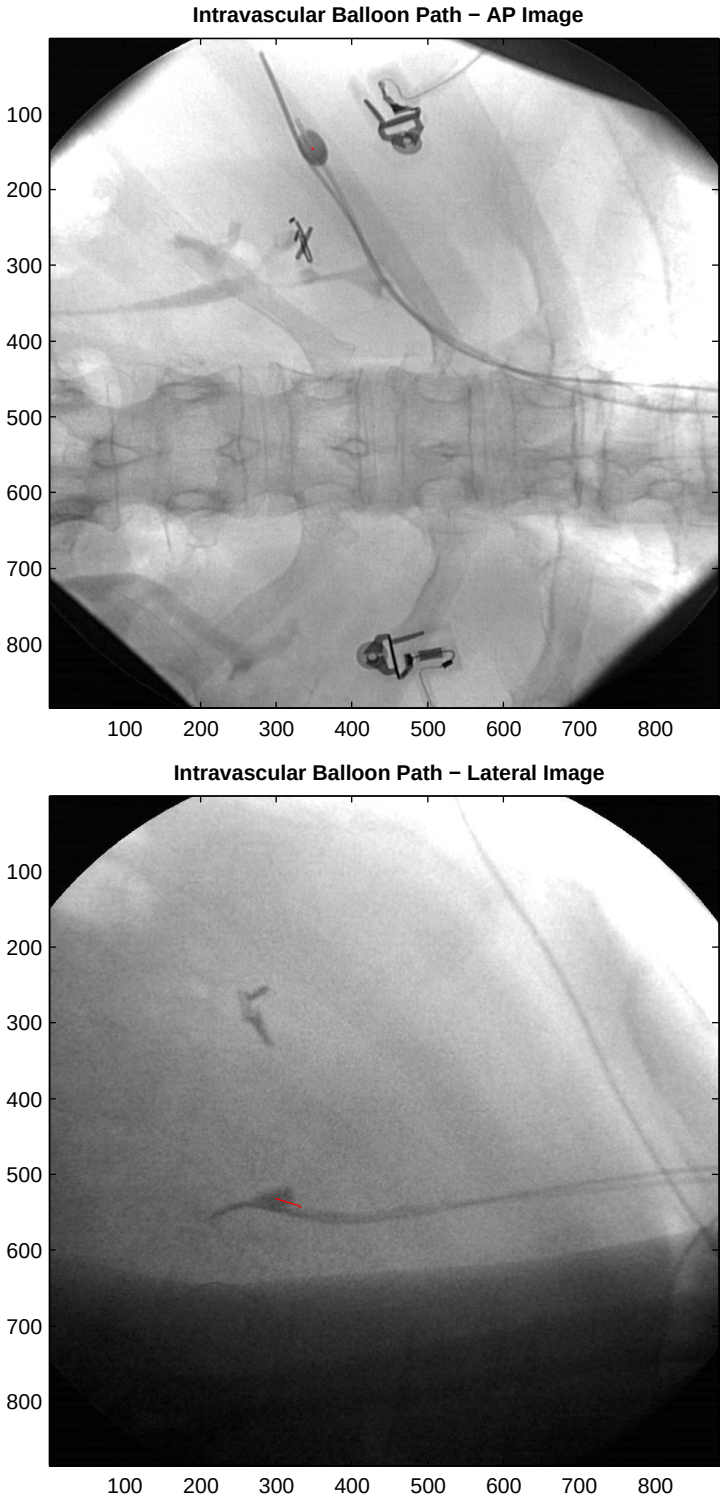


Figure 6.8: Balloon motion during TIPS011 breathing series 1. These images show the balloon motion during a respiratory cycle (PA top; LAT bottom).

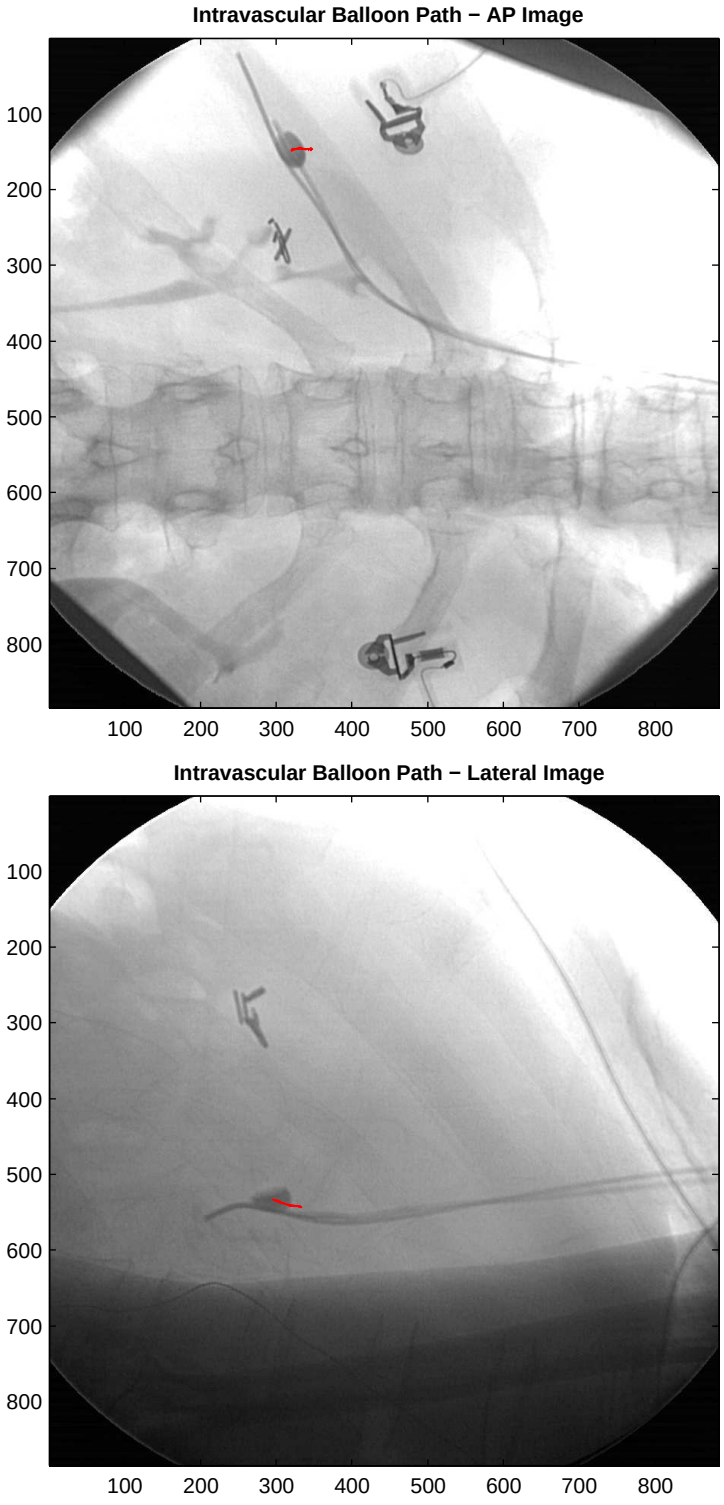


Figure 6.9: Balloon motion during TIPS011 breathing series 2. These images show the balloon motion during a respiratory cycle (PA top; LAT bottom).

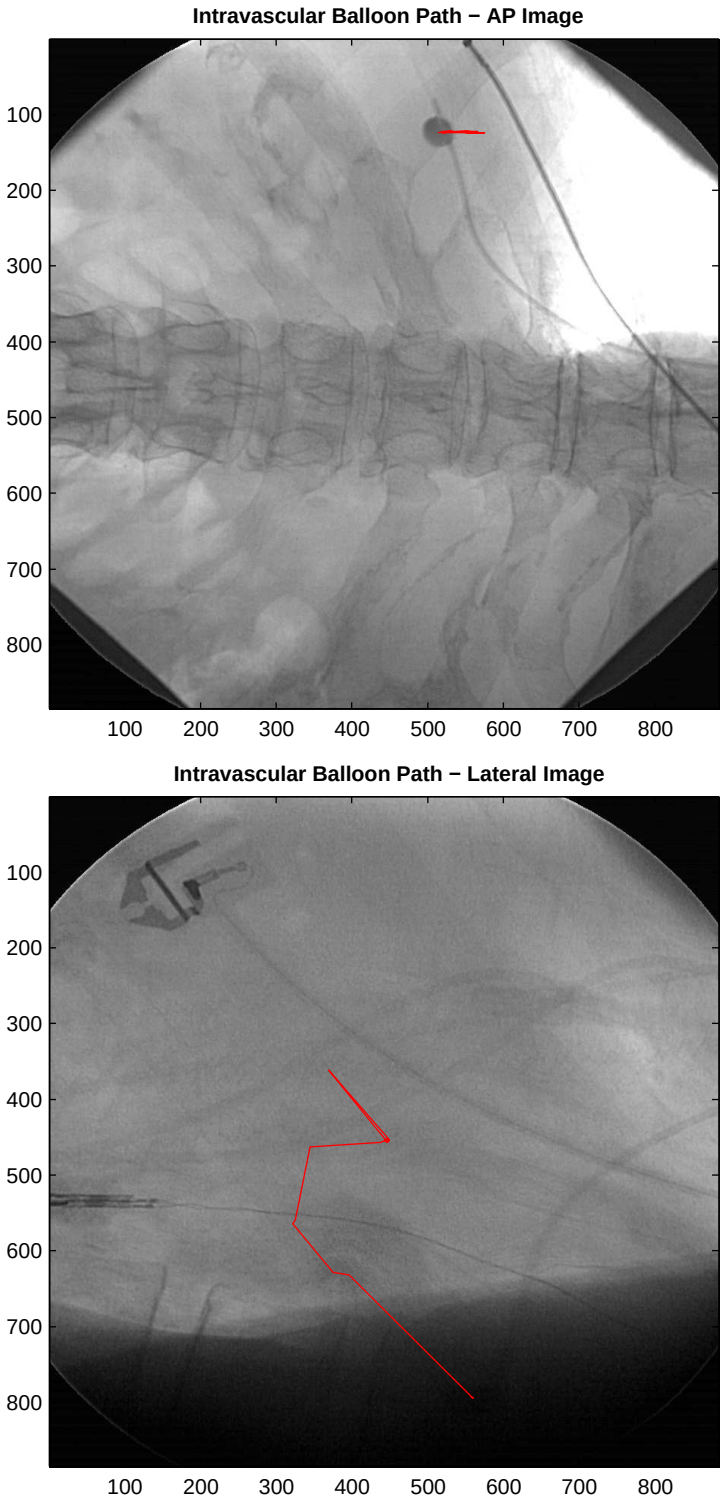


Figure 6.10: Balloon motion during TIPS012 breathing series. These images show the balloon motion during a respiratory cycle (PA top; LAT bottom). Note that the balloon was completely obscured by intervening tissue in most of the lateral series, so the tracker could not follow it during some frames. The balloon was not visible to human perception in these frames.

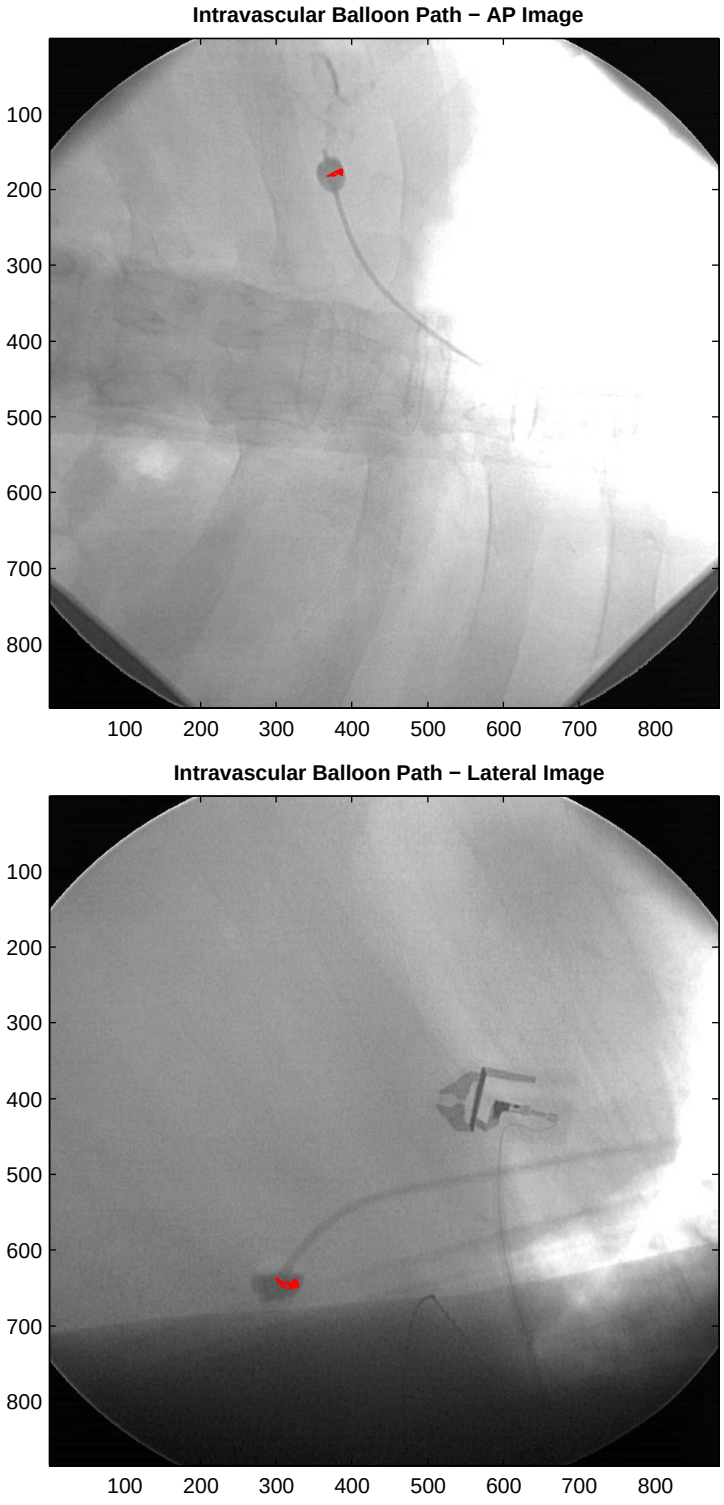


Figure 6.11: Balloon motion during TIPS013 breathing series. These images show the balloon motion during a respiratory cycle (PA top; LAT bottom).

Results of Balloon Tracking during Respiratory Motion

Item & Series		Range	Variance	Mean Change	Max Change
Rostro-caudal Translation					
(pixels)	*TIPS010-AP	-20.6 - 11.1	77.0	0.62	10.2
	*TIPS011-AP	-20.4 - 8.0	57.0	0.46	9.9
	*TIPS012-AP	-40.7 - 30.0	361.9	1.27	36.0
	*TIPS013-AP	-14.4 - 5.6	23.20	0.38	10.6
	*TIPS010-LAT	-32.0 - 18.8	180.3	1.09	19.2
	*TIPS011-LAT	-28.1 - 11.1	116.5	0.64	11.3
	*TIPS013-LAT	-21.2 - 9.1	67.21	0.63	12.3
Lateral Translation					
(pixels)	*TIPS010-AP	-3.6 - 2.9	2.26	0.16	3.3
	*TIPS011-AP	-2.7 - 2.6	2.14	0.22	3.3
	*TIPS012-AP	-2.6 - 1.32	0.77	0.11	2.6
	*TIPS013-AP	-5.4 - 4.4	9.28	0.64	7.9
Antero-posterior Translation					
(pixels)	*TIPS010-LAT	-6.1 - 7.2	13.64	0.44	8.0
	*TIPS011-LAT	-8.6 - 2.8	8.71	0.23	4.1
	*TIPS013-LAT	-9.7 - 5.7	14.11	0.98	11.7
Rotation in the Coronal Plane					
(degrees)	*TIPS010-AP	-10.94 - 24.23	88.55	0.840	24.20
	*TIPS011-AP	-1.20 - 1.06	0.16	0.114	1.32
	*TIPS012-AP	-5.91 - 52.20	33.30	1.64	49.47
	*TIPS013-AP	-1.56 - 1.21	0.189	0.174	2.14
Rotation in the Sagittal Plane					
(degrees)	*TIPS010-LAT	-1.97 - 6.47	5.30	0.300	4.67
	*TIPS011-LAT	-8.20 - 1.58	1.11	0.322	9.52
	*TIPS013-LAT	-3.00 - 1.65	0.921	0.279	2.30

Table 6.1: This table shows the range and variance of each element of the tracked motion as well as the average and maximum frame-to-frame change in each parameter.

6.3 Liver Motion during Needle Pressure

During the final phase of the procedure, the long TIPS needle is advanced into the hepatic vein, aimed, and forcefully pushed through the parenchyma of the liver towards the portal vein. Due to the substantial force involved in advancing the TIPS needle, one would expect a different magnitude and character of liver motion during these events. The subsequent data provides quantitive measurements of this motion.

Figure 6.12 includes graphs of the position and orientation variables over the four patient series, see Figure 6.15 for histograms of the displacements across all series, or see Figure 6.19 for a visual trace of one such balloon path over a needle pass. The motion of the liver was found to occur primarily in a craniocaudal direction (10mm aggregate range) with smaller variations in the AP and lateral position (3mm and 2mm ranges, respectively). The tracked location and orientation of the balloon in each frame corresponded visually with the true location of the balloon.

The TIPS012 series did not include images during a needle advancement phase, so it is excluded from the following analyses.

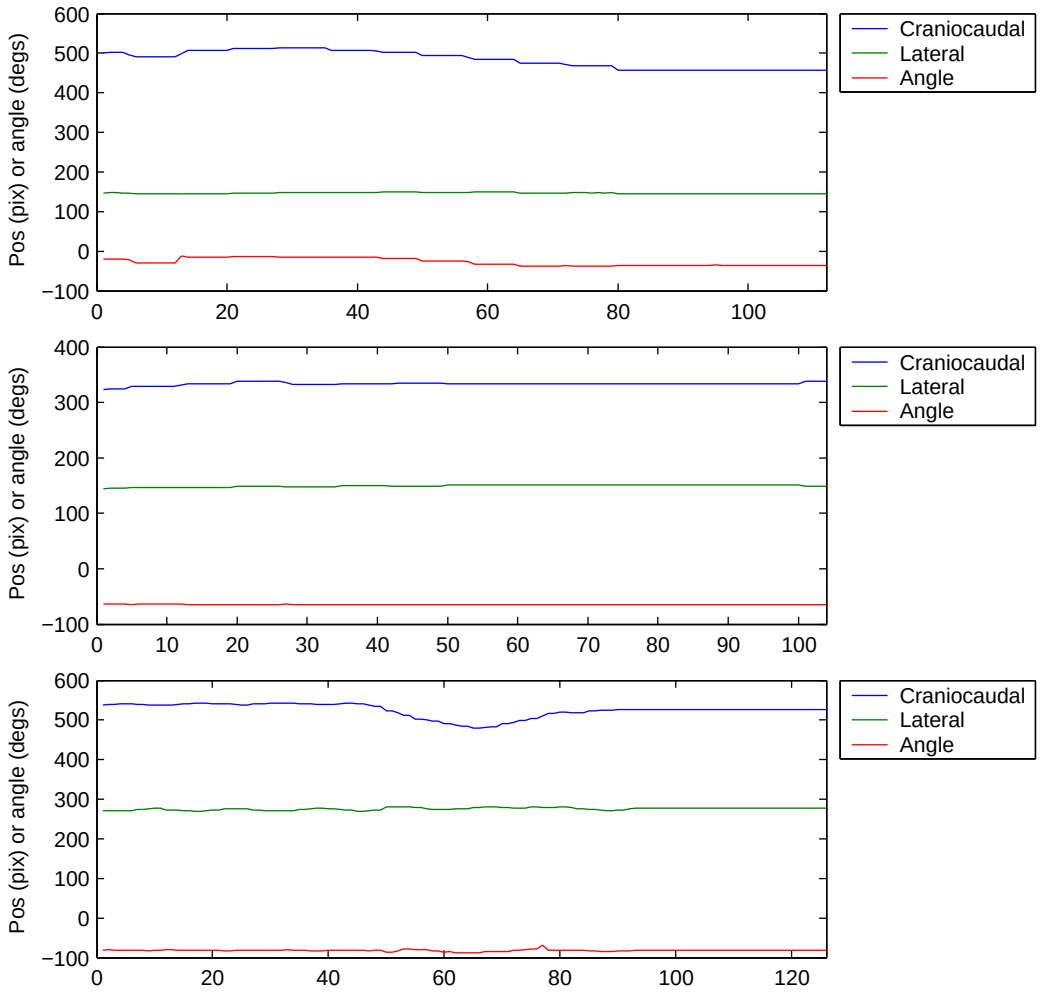


Figure 6.12: Motion of the intravascular balloon on PA images during needle pressure. The series shown from top to bottom are TIPS010, TIPS011, and TIPS013.

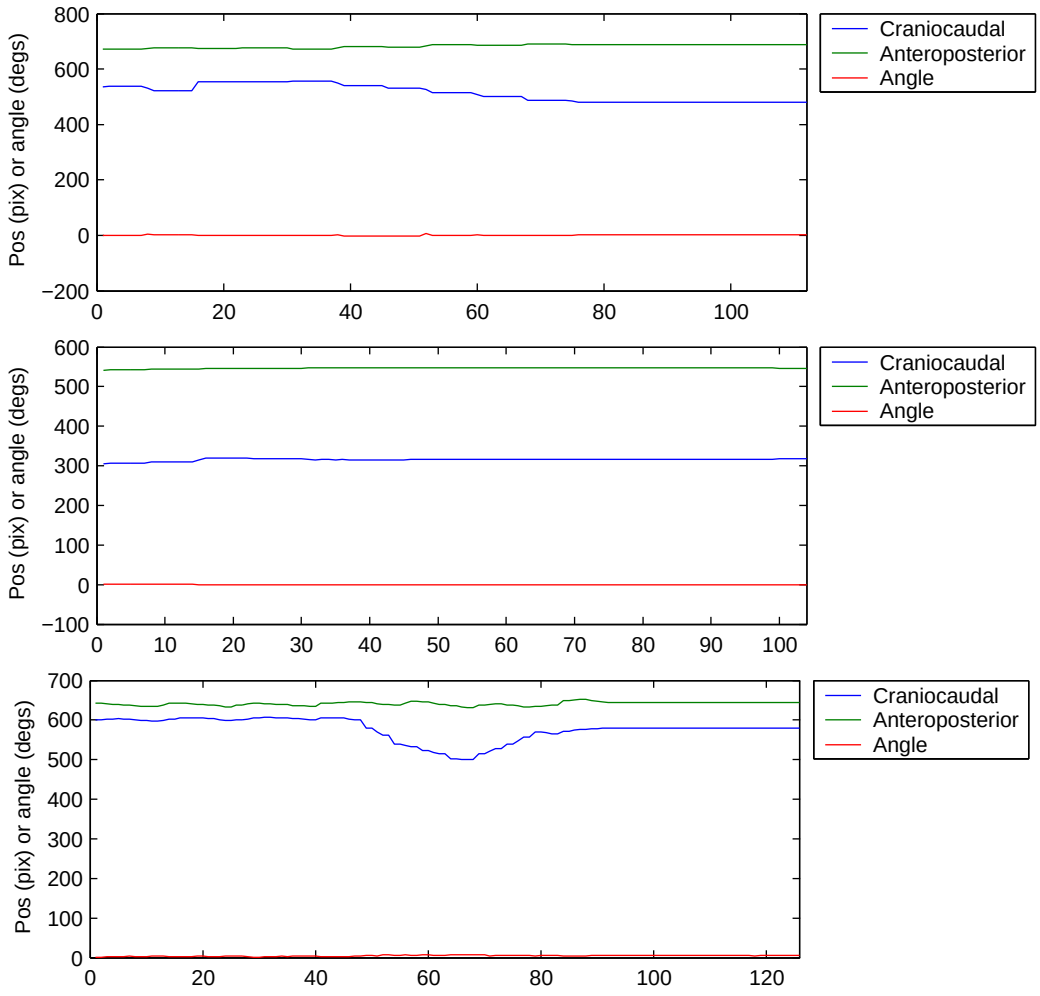


Figure 6.13: Motion of the intravascular balloon on lateral images during needle pressure. The series shown from top to bottom are TIPS010, TIPS011, and TIPS013.

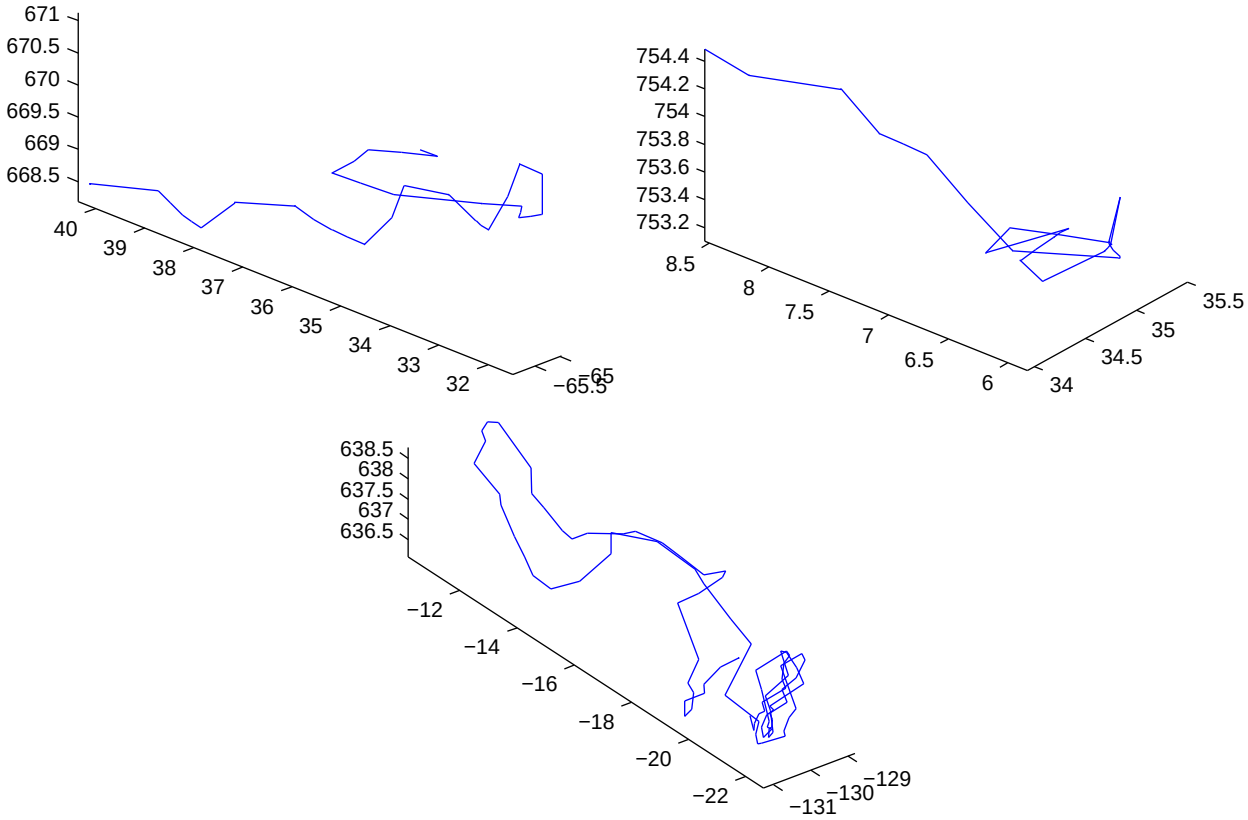


Figure 6.14: 3D motion of the intravascular balloon during needle pressure. The series shown from top to bottom are TIPS010, TIPS011, and TIPS013. Units are in millimeters.

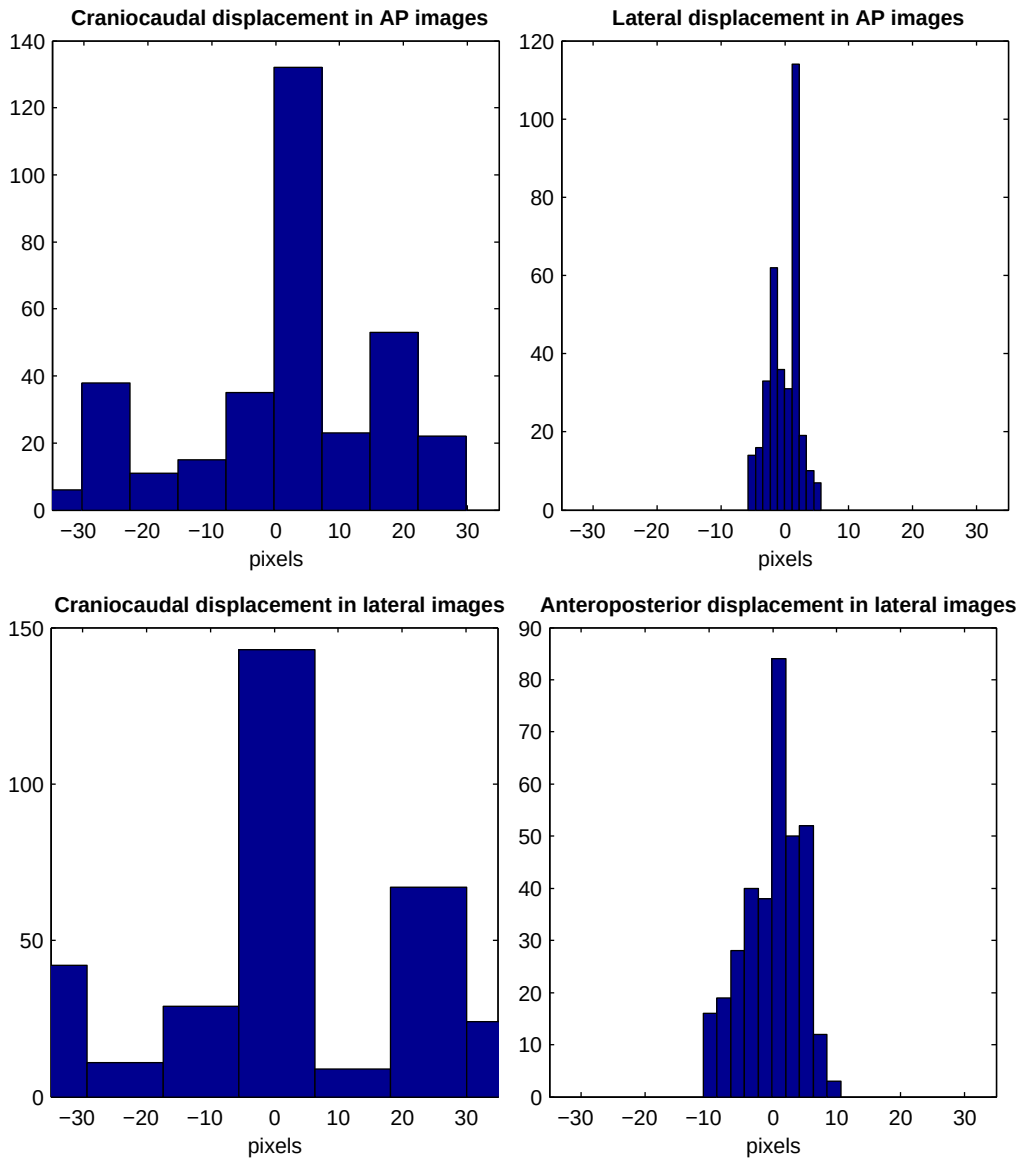


Figure 6.15: Distributions of displacements during needle pressure. The top row shows displacements in the PA images in the craniocaudal (left) and lateral (right) directions. The bottom row shows displacements in the lateral images in the craniocaudal (left) and PA (right) directions. Each histogram represents the aggregate results from all of the needle pressure series (TIPS010, TIPS011, and TIPS013).

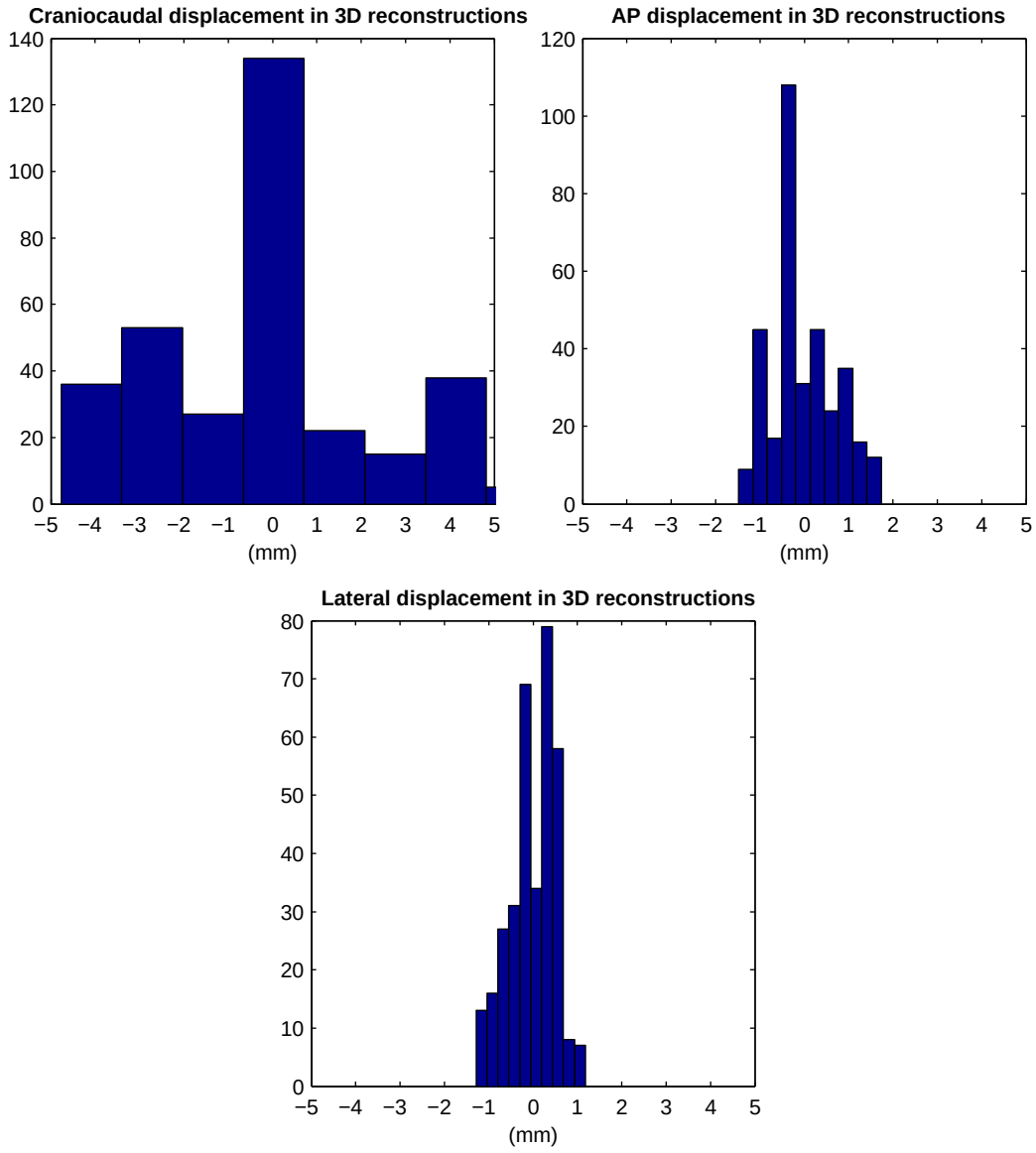


Figure 6.16: Distributions of 3D displacements during needle pressure. The graphs show 3D displacements in the craniocaudal (top left), PA (top right) and lateral (bottom right) directions. Each histogram represents the aggregate reconstruction results from all of the needle pressure series (TIPS010, TIPS011, and TIPS013).

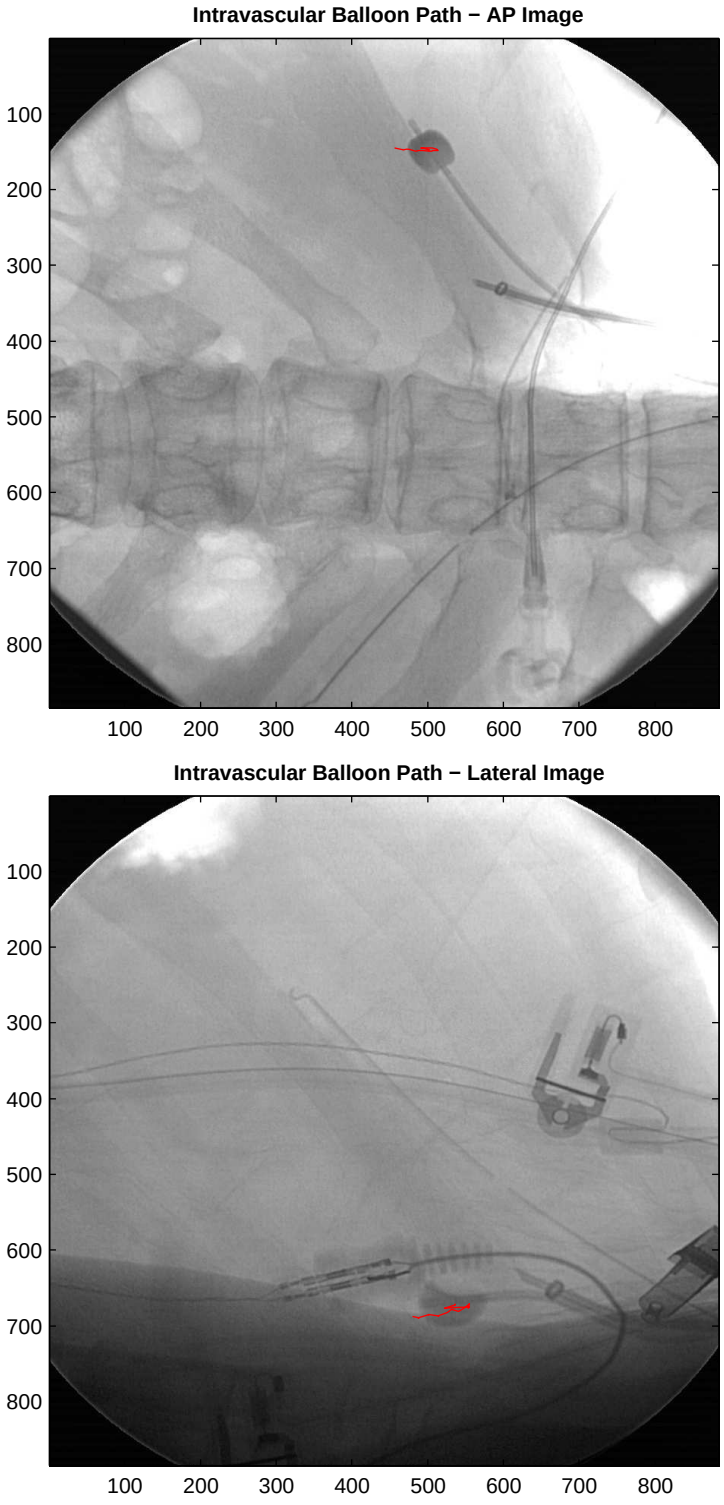


Figure 6.17: Balloon motion during TIPS010 needle pass series. These images show the balloon motion during a pass of the TIPS needle through the liver (PA top; LAT bottom).

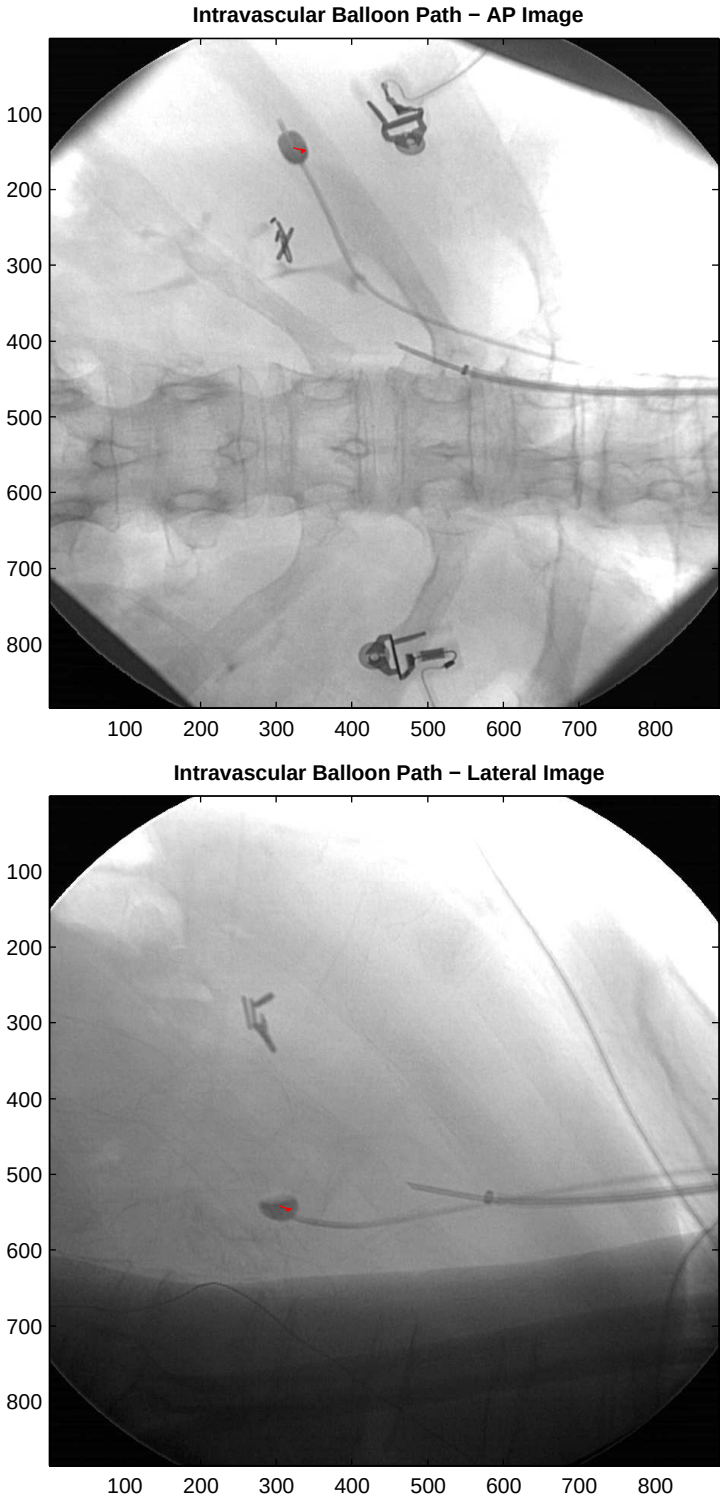


Figure 6.18: Balloon motion during TIPS011 needle pass series. These images show the balloon motion during a pass of the TIPS needle through the liver (PA top; LAT bottom).

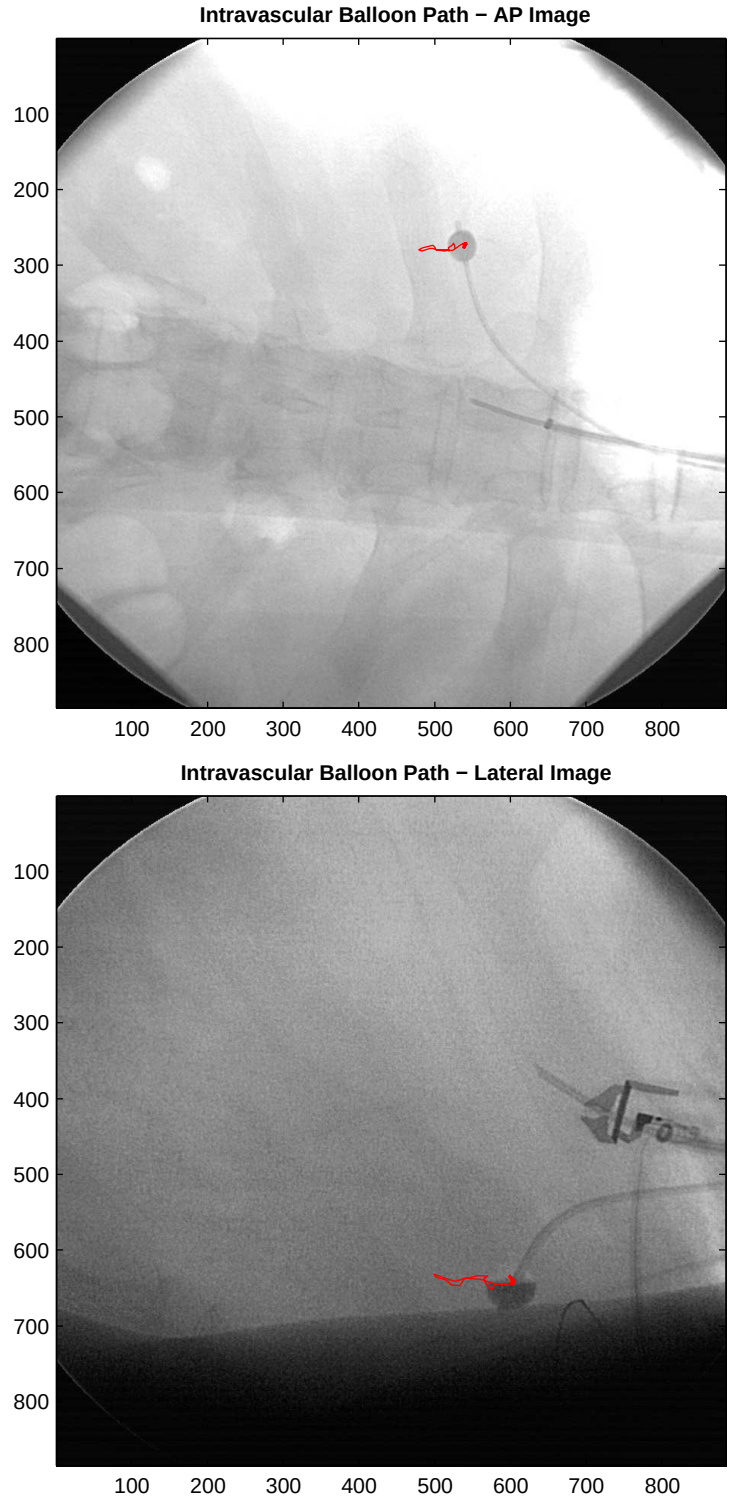


Figure 6.19: Balloon motion during TIPS013 needle pass series. These images show the balloon motion during a pass of the TIPS needle through the liver (PA top; LAT bottom).

Results of Balloon Tracking during Needle Pressure

Item & Series		Range	Variance	Mean Change	Max Change
Rostro-caudal Translation					
(pixels)	*TIPS010-AP	-27.6 - 29.8	470.9	0.86	11.4
	*TIPS011-AP	-10.7 - 4.8	7.95	0.28	5.3
	*TIPS013-AP	-45.3 - 17.7	300.2	1.11	10.8
	*TIPS010-LAT	-33.0 - 42.0	877.9	1.17	32.4
	*TIPS011-LAT	-10.7 - 3.3	11.26	0.24	4.7
	*TIPS013-LAT	-76.7 - 30.1	860.4	1.93	21.6
Lateral Translation					
(pixels)	*TIPS010-AP	-1.8 - 2.8	2.97	0.15	2.6
	*TIPS011-AP	-5.2 - 1.8	4.29	0.17	3.2
	*TIPS013-AP	-5.6 - 5.6	9.73	0.75	7.0
Antero-posterior Translation					
(pixels)	*TIPS010-LAT	-11.0 - 7.9	41.0	0.41	5.3
	*TIPS011-LAT	-5.1 - 1.4	2.67	0.10	2.0
	*TIPS013-LAT	-10.5 - 10.8	20.50	1.14	10.6
Rotation in the Coronal Plane					
(degrees)	*TIPS010-AP	-11.0 - 14.8	89.3	0.68	16.7
	*TIPS011-AP	-0.73 - 0.83	0.093	0.11	1.51
	*TIPS013-AP	-6.42 - 12.92	4.64	0.74	6.97
Rotation in the Sagittal Plane					
(degrees)	*TIPS010-LAT	-2.73 - 6.08	2.58	0.350	8.81
	*TIPS011-LAT	-0.65 - 1.78	0.52	0.12	2.42
	*TIPS013-LAT	-3.33 - 3.46	2.29	0.46	4.14

Table 6.2: This table shows the range and variance of each element of the tracked motion as well as the average and maximum frame-to-frame change in each parameter.

6.4 Conclusions

The results in this chapter suggest that there is substantive motion of the liver during both normal respiration and active needle advancement. The respiratory motion series show regular, rhythmic motion of the liver primarily in a craniocaudal direction (6mm range), but with small variations in the AP and lateral directions (3mm and 2mm ranges, respectively). The motion during needle pressure was variable and may be attributable to differences in the pressure being applied. In the first needle series, the liver appeared to slowly displace caudally and then hold its position. This may have been associated with slow, steady needle pressure. In the TIPS013 series, the liver showed a sudden, rapid caudal translation with subsequent return to its resting position. This may have represented a rapid jab of the needle. In all cases, motion occurred primarily in a craniocaudal direction (10mm aggregate range) with smaller movements in the AP and lateral directions (3mm and 2mm, respectively). These results are consistent with the movements that were anticipated based on the craniocaudal motion of the diaphragm and the anatomical constraints on the lateral, AP, and rotational motion of the liver. They are also consistent with prior published results on this issue [18, 56].

Part IV

Registration

The next part of this dissertation provides a theoretical development for a new method of 2D-3D registration. The motivation for this new method, its mathematical basis in biplane imaging, and the plan for validation will be discussed.

Chapter 7

2D-3D VASCULAR MODEL REGISTRATION

In order to display our pre-operative vascular model in proper alignment with our intraoperative data, it is necessary to estimate the relationship between the two coordinate frames. 2D-3D registration is a process that calculates the transformation that best projects a three-dimensional model to match one or more two-dimensional views. This chapter introduces a new approach to 2D-3D registration that makes use of the calibration work from Chapter 4. In particular, this approach optimizes over all available calibrated views simultaneously to estimate the best transformation, whereas prior methods have treated each view separately. A theoretical framework for this method is presented, as is a plan for validation.

7.1 Comparison of Existing Methods

There are three general approaches to 2D-3D registration in common practice. The first, intensity-based registration, uses a variety of measures of image intensity to compare acquired 2D data to calculated 2D projections. The second, feature-based registration, uses visible fiducials or landmarks to measure alignment. The final method, vessel-based registration, uses blood vessels extracted from 3D imaging and their corresponding 2D angiographic counterparts to perform registration. Each method will be discussed briefly below.

7.1.1 Intensity-Based Registration

Intensity-based registration uses the full host of pixel values in acquired 2D and calculated projections of 3D images to estimate the similarity of the two image sets. Graeme Penney et. al. present an excellent comparison of intensity-based registration methods in their 1998 IEEE-TMI article [74]. Some methods in this class seek to match the actual image intensities,

but many uses statistical measures of similarity that do not require identical intensity values between the images. For example, the mutual information method uses the frequency distributions of intensities in each image to measure the co-occurrence of values. When intensity values frequently occur together in the paired pixels, it is likely that the structures inducing those distributions are aligned. This allows, for example, the registration of 3D MR data to 2D fluoroscopic data despite the dramatic different in intensity assignments.

7.1.2 Feature-Based Registration

Feature-based registration relies on the extraction of fiducials or landmarks from 2D and 3D data for measures of registration accuracy. This approach is common in intraoperative tracking devices such as BrainLab^{TM}. In this approach, fiducials may be embedded prior to imaging (as in some radiation therapy planning approaches) or reliable landmarks such as bony prominences may be used. In the case of the BrainLab system, a laser pointer is used to establish point-by-point correspondence of a stable anatomical surface in 2D views. The optical registration between the reconstructed surface and the 3D imaging study is then determined.

When stable fiducials can be placed, these methods are generally considered to be the gold standard in registration. The challenges of placing stable fiducials (e.g., needing to screw markers into bony structures) limits the practicality of these methods.

7.1.3 Vessel-Based Registration

Another approach to 2D-3D registration is to use the vascular trees present in most anatomical structures as a basis for registration. This approach requires accurate extraction of 3D blood vessel models as well as extraction of vessels from 2D angiograms. This is arguably a subset of feature-based registration, but, due to the complexity that the registration of branching trees adds beyond that of point-to-point or simple surface-to-surface registration, it is treated separately in this work.

In Alan Liu's doctoral dissertation and related works, he describes a method of registration based upon the extraction and alignment of the midlines of discrete tubular objects [57–59]. His method begins with manually defined and associated pairs of 2D and 3D curves and optimizes

an iterative closest point metric to determine the optimal 3D rigid transform between the 3D model and camera coordinate systems. Importantly, Liu's method deals with only one view at a time. It also estimates both the model-to-world transformation and the camera calibration simultaneously. Because his approach serves as the basis for the new registration approach, his algorithm is summarized below.

The first step in Liu's approach is the definition or extraction of vessel curves from 2D and 3D data and the gross correspondence of segments between those data sources. For the 2D angiographic images, visible, unambiguous vessel segments were traced by hand in each view. The segments were manually correlated between 2D and 3D views. Vessels were initially traced manually from 3D data sets, but the advent of Aylward's methods yielded an automated approach to that problem [1].

The next step in the algorithm is the point-by-point association of elements in the 2D angiographic tracings with points on the projected 3D vessel segments. This association is made within each correlated vessel segment (from the prior step) using an iterative closest point approach. This method pairs the angiographic point $\tilde{X}^i_{angio}$ and the 3D point $\tilde{X}^i_w$ that projects most closely to it for the remainder of this iteration. Formally,

$$\tilde{X}^i_w = min_j \left\| \tilde{X}^i_{angio} - P\left(\tilde{\Pi}^j_w\right) \right\|^2 \tag{7.1}$$

where P is the function that projects 3D model points on to the angiographic view under analysis and $\tilde{\Pi}^j_w$ is a 3D vessel model point. Because the elements of P are the parameters being estimated by this algorithm, a reasonably good initial guess of the parameter values is needed as a starting point.

The next step in Liu's algorithm is the calculation of the error metric from the associated vessel points. This metric is a simple sum of error distances across all of the vessel segments, as follows.

$$\Phi = \sum_{i=1}^{n} \left\| \tilde{X}^i_{angio} - P\left(\tilde{X}^i_W\right) \right\|^2 \tag{7.2}$$

This error metric is optimized using the Levenberg-Marquardt variant of Newton's method with the intrinsic and extrinsic parameters of the camera system as the varying parameters. After each iteration of the Levenberg-Marquardt optimizer, the closest point relationships defined in Equation 7.1 are recalculated.

If additional views are involved (e.g., biplane systems), then each view is registered to the model separately. This approach requires the estimation of six parameters of a rigid registration and at least four intrinsic parameters of the camera projection matrix for each available view.

7.2 Calibrated, Stereoscopic Vessel-Based Registration

In this section, I will extend the methods presented in Liu's dissertation [59] to accommodate simultaneous registration of a three-dimensional model to two or more projection views on a calibrated stereoscopic imaging system. For ease of reference, this method will be referred to as the CAlibrated Stereoscopic TubulE Registration (CASTER) algorithm.

In single-view registration, the depth of an object is generally determined with less certainty than the other components of its position since the effects of perspective scaling are weak as compared to those of orthogonal displacement. By using two views simultaneously for calibration, especially two orthogonal views, one can rely upon non-perspective cues for determination of all aspects of the object-to-camera-frame transformation. In addition, the use of precalibrated cameras reduces the number of degrees of freedom (DOF) in the registration from twenty-two (22) to six (6) because the calibration provides intrinsic parameters for both cameras ($5x2 = 10$ DOF) and the intercamera extrinsic parameters (6 DOF), leaving only the object-to-camera-frame extrinsic parameters (6 DOF) to be determined.

7.2.1 Weighting of Registration Errors Across Points and Cameras

Let us first consider the error between a point on an angiogram and the nearest corresponding point from the projected model of that blood vessel. If we have a point on the vessel $\tilde{X}_w$, expressed in world coordinates, then the projection of that point into camera coordinates is governed by the standard projection equations:

$$\tilde{X}_{cam} = M_{int}\, M_{w \rightarrow cam}\, \tilde{X}_w \qquad \text{and, in image coordinates,} \tag{7.3}$$

$$x_{im} = \frac{x_{cam}}{z_{cam}} \tag{7.4}$$

$$y_{im} = \frac{y_{cam}}{z_{cam}} \tag{7.5}$$

where M_{int} is the internal projection matrix for the camera and $M_{w \to cam}$ is the external transformation matrix from world to camera coordinates. The two-dimensional reprojection error between the angiographic and projected point is thus:

$$ERR_{im} = \left\| \tilde{X}_{im} - \tilde{X}_{angio} \right\|^2 \tag{7.6}$$

Let us next consider the magnitude of the 2D reprojection error for a give three-dimension registration error. Define $\tilde{X}_w^\alpha$ to be the true world coordinates of a point on a vessel. Let $\tilde{X}_w^\beta$ be the world coordinates of the same point under the current registration transformation. Let k be the magnitude of the three dimensional positional error. For ease of analysis, we will restrict this error to lie within a plane parallel to the image plane. We then have:

$$\tilde{X}_w^\alpha = \begin{bmatrix} x_w^\alpha \\ y_w^\alpha \\ z_w^\alpha \end{bmatrix} \quad \text{and} \tag{7.7}$$

$$\tilde{X}_w^\beta = \begin{bmatrix} x_w^\alpha + k_x \\ y_w^\alpha + k_y \\ z_w^\alpha \end{bmatrix} \quad \text{such that} \tag{7.8}$$

$$\left\| \tilde{X}_w^\alpha - \tilde{X}_w^\beta \right\| = k \tag{7.9}$$

$$\Rightarrow k = \sqrt{k_x^2 + k_y^2} \tag{7.10}$$

On projection, we have:

$$\tilde{X}^{\alpha}_{cam} = M_{int}\ M_{w\to cam}\ \tilde{X}^{\alpha}_{w} \tag{7.11}$$

$$\tilde{X}^{\alpha}_{cam} = \begin{bmatrix} f_x & s & oc_x \\ 0 & f_y & oc_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & Tz \end{bmatrix} \begin{bmatrix} x^{\alpha}_{w} \\ y^{\alpha}_{w} \\ z^{\alpha}_{w} \\ 1 \end{bmatrix} \tag{7.12}$$

$$\tilde{X}^{\alpha}_{cam} = \begin{bmatrix} f_x & s & oc_x \\ 0 & f_y & oc_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^{\alpha}_{w} \\ y^{\alpha}_{w} \\ z^{\alpha}_{w} + T_z \end{bmatrix} \tag{7.13}$$

and, by components,

$$x^{\alpha}_{cam} = f_x\ x^{\alpha}_{w} + s\ y^{\alpha}_{w} + oc_x\,(z^{\alpha}_{w} + T_z) \tag{7.14}$$

$$y^{\alpha}_{cam} = f_y\ y^{\alpha}_{w} + oc_y\,(z^{\alpha}_{w} + T_z) \tag{7.15}$$

$$z^{\alpha}_{cam} = z^{\alpha}_{w} + T_z \tag{7.16}$$

The final image coordinates are therefore

$$x^{\alpha}_{im} = \frac{f_x\ x^{\alpha}_{w} + s\ y^{\alpha}_{w}}{z^{\alpha}_{w} + T_z} + oc_x \tag{7.17}$$

$$y^{\alpha}_{im} = \frac{f_y\ y^{\alpha}_{w}}{z^{\alpha}_{w} + T_z} + oc_y \tag{7.18}$$

The derivation for the projected model point, X^{β}_{w}, yields a similar result:

$$x^{\beta}_{im} = \frac{f_x\,(x^{\alpha}_{w} + k_x) + s\ y^{\alpha}_{w}}{z^{\alpha}_{w} + T_z} + oc_x \tag{7.19}$$

$$y^{\beta}_{im} = \frac{f_y\,(y^{\alpha}_{w} + k_y)}{z^{\alpha}_{w} + T_z} + oc_y \tag{7.20}$$

The 2D reprojection error between the ideal vessel point, $\tilde{X}^{\alpha}_{w}$ and the corresponding registered model point $\tilde{X}^{\beta}_{w}$ can now be calculated as:

$$\left\|\tilde{X}_{im}^{\alpha} - \tilde{X}_{im}^{\beta}\right\| = \sqrt{\left(\frac{f_x\ x_w^{\alpha} + s\ y_w^{\alpha}}{z_w^{\alpha} + T_z} + oc_x - \frac{f_x\left(x_w^{\alpha} + k_x\right) + s\ y_w^{\alpha}}{z_w^{\alpha} + T_z} - oc_x\right)^2 + \left(\frac{f_y\ y_w^{\alpha}}{z_w^{\alpha} + T_z} + oc_y - \frac{f_y\left(y_w^{\alpha} + k_y\right)}{z_w^{\alpha} + T_z} - oc_y\right)^2} \tag{7.21}$$

$$= \sqrt{\left(\frac{f_x\ x_w^{\alpha} + s\ y_w^{\alpha} - f_x\left(x_w^{\alpha} + k_x\right) - s\ y_w^{\alpha}}{z_w^{\alpha} + T_z}\right)^2 + \left(\frac{f_y\ y_w^{\alpha} - f_y\left(y_w^{\alpha} + k_y\right)}{z_w^{\alpha} + T_z}\right)^2}$$

$$\left\|\tilde{X}_{im}^{\alpha} - \tilde{X}_{im}^{\beta}\right\| = \sqrt{\left(\frac{f_x\ k_x}{z_w^{\alpha} + T_z}\right)^2 + \left(\frac{f_y k_y}{z_w^{\alpha} + T_z}\right)^2} \tag{7.22}$$

The last equation shows that the measured reprojection error is not only determined by the actual three-dimensional position error, but is also influenced by the camera's focal length, aspect ratio, and the distance from the point to the camera. These dependencies are undesirable in the error metric because they induce biases in the optimization process. These biases include weighting errors closer to the camera more heavily than errors in registration of distant points and magnifying errors in cameras with longer focal lengths (narrower fields of view).

These biases may be reduced or eliminated by rescaling the measured reprojection errors prior to summation in the metric. Let the measured 2D reprojection error for our point pair be expressed as a 2D vector $\begin{bmatrix} x_{err}^{\alpha} & y_{err}^{\alpha} \end{bmatrix}$. We can then define a normalized error metric as follows:

$$x_{NormErr}^{\alpha} = x_{err}^{\alpha}\left(\frac{z_w^{\alpha} + T_z}{f_x}\right) \tag{7.23}$$

$$y_{NormErr}^{\alpha} = y_{err}^{\alpha}\left(\frac{z_w^{\alpha} + T_z}{f_y}\right) \tag{7.24}$$

$$\tag{7.25}$$

Combining these components yields

$$Err_{norm}^{\alpha} = \sqrt{\left[x_{err}^{\alpha}\left(\frac{z_w^{\alpha} + T_z}{f_x}\right)\right]^2 + \left[y_{err}^{\alpha}\left(\frac{z_w^{\alpha} + T_z}{f_y}\right)\right]^2} \tag{7.26}$$

and by substitution

$$Err_{norm}^{\alpha} = \sqrt{\left[\left(\frac{f_x\ k_x}{z_w^{\alpha} + T_z}\right)\left(\frac{z_w^{\alpha} + T_z}{f_x}\right)\right]^2 + \left[\left(\frac{f_y k_y}{z_w^{\alpha} + T_z}\right)\left(\frac{z_w^{\alpha} + T_z}{f_y}\right)\right]^2} \tag{7.27}$$

$$Err_{norm}^{\alpha} = \sqrt{k_x^2 + k_y^2} \tag{7.28}$$

$$Err_{norm}^{\alpha} = k \tag{7.29}$$

The final equation shows that reweighting the measured component errors as in Equation 7.26 yields a new error metric that is solely dependent upon the distance between the actual and estimated point locations. This approach forms the basis for a new, unbiased approach to joint 2D-3D registration across two or more views.

7.2.2 Weighted Error Metric for One or More Views

Having defined an improved error metric in the prior section, one must now consider the application of such a metric across one or more views. Equation 7.2 provided the basic metric for registration of a single view. One can expand this metric to include additional views as follows:

$$\Phi = \sum_{j=1}^{m} \sum_{i=1}^{n_j} \left\| \tilde{X}_{angio}^{ij} - P^j \left(\tilde{X}_w^{ij} \right) \right\|^2 \tag{7.30}$$

where m is the number of views, n_j is the number of visible points in view j, $\tilde{X}_{angio}^{ij}$ is the location of the i^{th} point on the j^{th} angiographic view, P_j is the function that projects model points on to the image plane of the j^{th} view, and $\tilde{X}_w^{ij}$ is the location of the nearest neighbor to the i^{th} angiographic point in the j^{th} view expressed in 3D world coordinates. We may then modify this metric to incorporate the error weighting method from the prior view, as follows:

$$\Phi_{norm} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} \left(\left[x_{err}^{ij} \left(\frac{z_w^{ij} + T_z^j}{f_x^j} \right) \right]^2 + \left[y_{err}^{ij} \left(\frac{z_w^{ij} + T_z^j}{f_y^j} \right) \right]^2 \right) \quad \text{where} \tag{7.31}$$

$$x_{err}^{ij} = x_{angio}^{ij} - P_x^j \left(\tilde{X}_w^{ij} \right) \tag{7.32}$$

$$x_{err}^{ij} = x_{angio}^{ij} - \left(\frac{M_{int_x}^j M_{w \to j}^j \tilde{X}_w^{ij}}{M_{int_z}^j M_{w \to j}^j \tilde{X}_w^{ij}} \right) \quad \text{and} \tag{7.33}$$

$$y_{err}^{ij} = y_{angio}^{ij} - P_y^j \left(\tilde{X}_w^{ij} \right) \tag{7.34}$$

$$y_{err}^{ij} = y_{angio}^{ij} - \left(\frac{M_{int_y}^j M_{w \to j}^j \tilde{X}_w^{ij}}{M_{int_z}^j M_{w \to j}^j \tilde{X}_w^{ij}} \right) \tag{7.35}$$

In these equations, x_{err}^{ij} and y_{err}^{ij} represent the x and y components of the 2D error between the i^{th} point in the j^{th} angiographic view and the nearest projected model point. $M_{int_x}^j$ represents the x-axis component of the internal projection matrix for the j^{th} view, and $M_{w \to j}^j$ represents the matrix that transforms from the world coordinate system to the coordinates of the j^{th} camera.

7.2.3 Optimization of Registration Using Weighted Error Metric

The goal of the 2D-3D registration process is to determine the rigid transformation matrix $M_{model \rightarrow w}$ that best relates the coordinate frame of the object model to that of the world system. Formally, one wishes to find the euclidean transformation matrix $M_{model \rightarrow w}$ in $\tilde{X}_w^{ij} = M_{model \rightarrow w} \tilde{X}_{model}^{ij}$ that minimizes the error metric Φ_{norm} in Equation 7.31. This can be accomplished via iterative optimization of Φ_{norm} over the six parameters of $M_{mod \rightarrow w}$.

7.3 Approach to Validation

While validation data for the aforementioned algorithm will not be presented within the scope of this work, a brief discussion of the planned approach to validation may be of value to the reader. For this purpose, an outline of the validation methods is discussed below.

In order to evaluate the performance of the CASTER algorithm, a series of six physical models of portal venous trees were constructed and imaged. The models were designed to allow construction of accurate 3D vessels models from CT images, provide distinct vessel-like projections under fluoroscopy, and to provide an independent estimate of the ideal transformation from model to world coordinates.

The models were made using heavy pottery clay; they mimic the structure of the portal venous tree. The x-ray absorption characteristics of the heavy clay are similar to that of a large contrasted vessel. The models image well using fluoroscopy; Figure 7.1 shows simultaneous PA and lateral views. CT and optical images of one of the phantoms are shown in Figure 7.2. The models were stabilized using low-density rigid foam, and brass bearings were added for use as quantifiable reference fiducials (dark circles in CT and fluoroscopic images).

Validation will involve several stages of image analysis and processing, as follows:

- Calibration - The internal and external parameters of the fluoroscopic imaging system will be determined using the methods described in Chapter 4.

- Ground Truth Registration - The embedded fiducials will be extracted from the fluoroscopic and CT images and used to measure the correct transformation from CT coordinates to the fluoroscopic world coordinate frame.

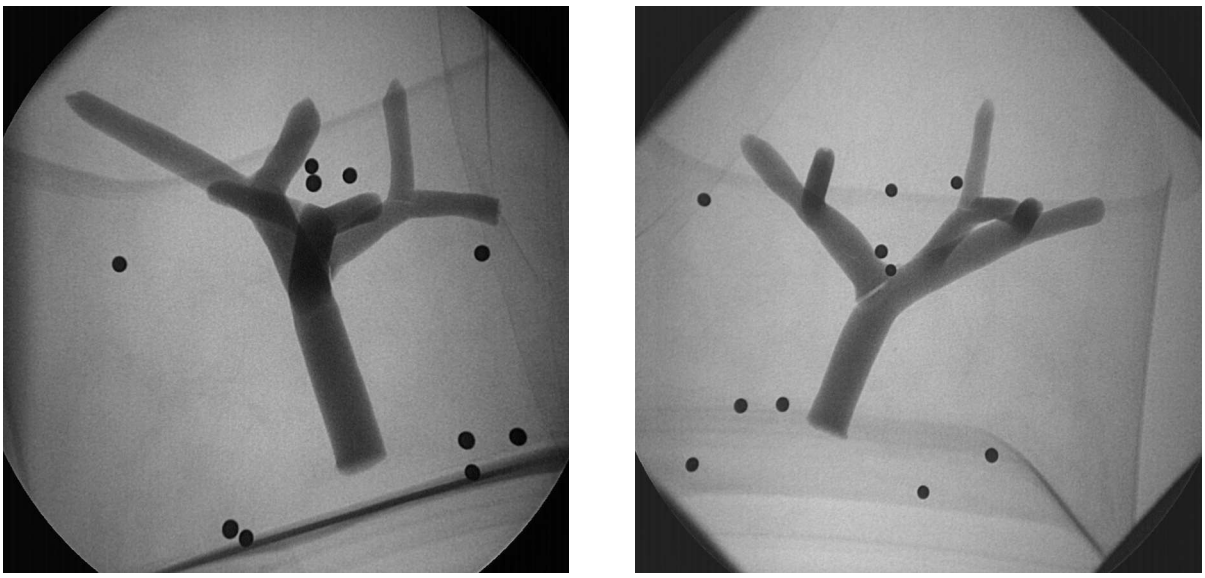


Figure 7.1: Fluoroscopic images of one of the vessel phantoms. The PA (left) and lateral (right) views show tubular structures similar that mimic the patterns of the portal venous system. Embedded fiducials (spherical bearings) are also visible as small circles throughout the images.

Figure 7.2: CT and optical images of one of the vessel phantoms. The CT image (left) shows segments of the vessel branches as well as slices of the embedded fiducials. The optical image (right) shows the branching structure of the model and its relative size.

- 3D Vessel Model - Using Aylward's vessel extraction methods, a 3D model of each vessel phantom will be constructed from CT data.

- Fluoroscopic Vessel Tracings - Vessel segments will be manually traced and related between fluoroscopic views. Correspondence with 3D vessel model segments will be established manually as well.

- Comparison of Standard Method and CASTER Method - 2D-3D registrations will be calculated using Liu's approach and the new multi-view weighted error approach. These registrations will be compared to the ground truth transformation calculated from the fiducials.

7.4 Conclusion

This chapter has introduced a new approach to 2D-3D registration that uses the redundancy of biplane views and the accurate calibration methods of Chapter 4 to advance the state of the art. This new approach involves optimizing six parameters using all available views simultaneously, whereas prior methods have involved the optimization of ten or more parameters in each view separately. A novel validation process is underway to determine the potential benefits of this new approach to 2D-3D registration.

Part V

Integration and Conclusion

The final part of this dissertation will integrate and summarize the material from the prior chapters. Opportunities for future research will also be discussed.

Chapter 8

SUMMARY OF INNOVATIONS

The purpose of this dissertation was to advance several of the methods critical to precision tracking and guidance for intra- and perivascular procedures. The most important innovations in this work are in the calibration of fluoroscopic systems. These innovations include: a new calibration phantom that allows simultaneous calibration and distortion correction from two views, making accurate intraoperative calibration more practical; an integrated bundle adjustment method that allows simultaneous calibration and distortion correction while fully exploiting the redundancy of biplane imaging; and a rigorous experimental study of accuracy, demonstrating mean errors of 0.44mm. This work substantially advances prior published work showing theoretical mean errors of 1-2mm [29, 64].

The next part of this work discussed intraoperative target tracking in the context of following moving anatomical targets during a procedure. The main innovations of this part includes the use of an intravascular balloon catheter as a tracking fiducial and a method of identifying and tracking a moving target using a maximum likelihood approach with a diverse set of feature likelihood distributions. This method provides a robust means of finding the desired balloon fiducial from among a field of anatomical and instrumentation distractors. This work is distinct from prior work using surface markers or intraoperative ultrasound for tracking and, as such, may represent a new branch in intraoperative tracking research.

The tracking methods in this dissertation also provided a view of the motion of the liver during the TIPS procedure. In analyses of intraoperative data from four patients, respiratory motion ranged over 6mm in the craniocaudal direction, 3mm anteroposteriorly, and 2mm laterally. The motion during needle pressure followed a similar pattern; motion ranges were 10mm, 3mm, and 2mm in the craniocaudal, AP, and lateral directions respectively. These

results are consistent with the anatomically expected motion of the liver.

The final innovation in this work is a method of 2D-3D registration that explicitly exploits the redundancy of biplane imaging geometries. The new method includes an optimization approach that optimizes 2D reprojection error in both views simultaneously. This method also used a new means of weighting these errors based on the depth of the corresponding feature. While experimental evaluation of this method is ongoing, the theoretical development of the new algorithm represents a substantial advance over prior single-view registration methods.

Overall, this work provides a meaningful basis for the development of a clinically useful intraoperative tracking and guidance system. The calibration methods provide a better than four-fold improvement over prior methods, lowering the overall error in subsequent tracking and registration steps. The new approach to organ tracking has, in our experimental series, been able to robustly track our new intravascular fiducial despite numerous nearby distractor objects. This is critical to providing reliable intraoperative guidance updates. The final piece of this work, registration, makes full use of the biplane imaging system to align models to patients. When combined with high-fidelity vessel modeling and good display systems, these results provide the basis for a complete intraoperative guidance system based upon biplane fluoroscopy.

Chapter 9

FUTURE WORK

One of the most challenging aspects of this dissertation was knowing when to defer an avenue of research for future work. This chapter identifies and briefly discusses several of these deferred issues.

9.1 Calibration

The primary need for the next generation of calibration methods is a more accurately manufactured calibration phantom. In particular, the next generation of phantom absolutely has to have rigid mounting holes of known depth for the fiducial bearings; allowing the bearings to rest and move along a drill track creates many problems. As demonstrated by the nearly 50% reduction in reconstruction error associated with mathematical correction of model errors, manufacturing problems can have a very significant impact on overall results. The use of thicker plexiglass planes would allow for precisely placed drill wells. On the downside, this approach would significantly increase the cost of making such phantoms.

Another issue in calibration is the increasing popularity of flat-panel x-ray detectors. As this alternative to the image intensifier tube progressively takes over the market, a different set of calibration and distortion correction concerns will need to be addressed. It may be become important, for example, to explicitly model the tilt of the detector plane relative to the optical axis. It is likely that these systems will have much less need for high-order distortion correction than is necessary in IIT systems.

A final area of personal curiosity in calibration was the problem of identifying the true projected center of the spherical fiducials. Although all spheres project to conics (ellipses except in degenerate cases), the 3D center of that sphere does not generally project to the

2D center of the ellipse. Instead, the true projected center lies somewhere along the major axis of the ellipse (away from the optical center of the system). While the magnitude of this effect is sub-pixel in the current system, it does introduce a systematic radial bias into the measurements that may have larger cumulative effects.

One long-term goal of this research is the development of a simple, unobtrusive calibration process that would allow for dynamic calibration as the biplane view was changed. The use of a distinct calibration object is an obstacle to such a process. One possibility would be to adapt Fencil & Metz's methods for calibration using unstructured scene elements to use semi-structured elements. If one were to place a sufficient number of internally known targets, such as short rods or triangles of known size with fiducials on each end/vertex, into the field of view (on the patient's skin), then it may be possible to conduct dynamic, robust bundle adjustment without the need for patient repositioning or user intervention. The error characteristics of such an approach will have to be explored, as will the impact on the visibility of the medically important features in the scene.

9.2 Tracking

The first opportunity for improvement in the tracking methods lies in the standardization of the intraoperative balloon placement process. In the four data series presented herein, each series appeared to use a different size balloon, different anatomical placement, and varying levels of filling of the balloon with contrast. It will be important to be more consistent in each of these areas so that the tracker may be better optimized to a predictable set of parameters.

The next need in tracking is the evaluation of these methods using additional patient data. It would be desirable to generate a database of balloon and distractor feature parameters that could be used to tune the probability distributions for the maximum likelihood method. Additional data would also allow us to quantify the variability of liver motion during respiration and needle advancement in a more substantial population.

Additional areas of future tracking work are discussed in §5.4.

9.3 Registration

Validation is the first concern with the CASTER registration algorithm. Analysis of the vessel models is underway currently and should yield results within the next few months.

There are several areas for improvement in this approach to registration. The current use of manual midline segmentation and correlation between views is tedious and prone to error. Automated methods should be developed to accomplish each of these tasks. The use of distinct vessel segments is itself somewhat problematic; it may be desirable to shift to identifying a pool of possible vessel points and using a combination of epipolar constraints and geometric correlation with the 3D vessel model to establish the relationship between the 2D and 3D structures.

In the longer view, it may be possible to integrate the construction of the 2D model elements, correlation with 3D model elements, and estimation of optimal registration into a single iterative process. One could use the structure suggested by the 3D model and the estimated registration to derive more robust 2D vessel trees and to improve correlation with the 3D model. Similarly, the 3D model could be refined as the process progressed.

A final area of consideration in the CASTER algorithm is the effect of the weighting scheme on the distribution of errors in the scene. While the weighting method has been shown to provide a less biased metric of actual 3D registration error, the weighting also affects the magnitude of the errors associated with each point. If one assumes, for example, that the error in feature localization in 2D images has a constant distribution across the image, then that distribution will be magnified by the reweighting of distant points and constricted for nearer points. This skewing of error distributions may induce a slightly different bias into the optimization method. It may be possible to adjust the weights by a small amount to account for the uncertainty associated with these changes in error distributions.

BIBLIOGRAPHY

[1] S. Aylward. *Gaussian Goodness of Fit Cores*. Phd thesis, University of North Carolina at Chapel Hill, 1997.

[2] S. Aylward and E. Bullitt. Initialization, noise, singularities, and scale in height-ridge traversal for tubular object centerline extraction. *IEEE Transactions on Medical Imaging*, pages 61–75, 2002.

[3] S. Aylward, E. Bullitt, S. Pizer, and D. Eberly. Intensity ridge and widths for tubular object segmentation and description. *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 131–8, 1996. (03) Dept. of Comput. Sci., North Carolina Univ., Chapel Hill, NC, USA.

[4] Ronald T. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[5] M. Bajura, H. Fuchs, and R. Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *Computer Graphics, Proceedings of SIGGRAPH 1992*, 26(2):203–210, 1992.

[6] R. Banares, M. Casado, J. M. Rodriguez-Laiz, F. Camunez, A. Matilla, A. Echenagusia, G. Simo, B. Piqueras, G. Clemente, and E. Cos. Urgent transjugular intrahepatic portosystemic shunt for control of acute variceal bleeding. *American Journal of Gastroenterology*, 93(1):75–9, 1998.

[7] Y. Bar-Shalom, H.M. Shertukde, and K.R. Pattipati. Use of measurements from an imaging sensor for precision target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 25(6):863–72, 1989. (03) Dept. of Electr. & Syst. Eng., Connecticut Univ., Storrs, CT, USA (39) IEARAX.

[8] P. M. Black, T. Moriarty, E. Alexander, P. Stieg, E. J. Woodard, P. L. Gleason, C. H. Martin, R. Kikinis, R. B. Schwartz, and F. A. Jolesz. The development and implementation of intraoperative mri and its neurosurgical applications. *Neurosurgery*, 41:831–42, 1997.

[9] P. Bonard, G. Tremblay, A. R. LeBlanc, M. Bertrand, and F. A. Roberge. A new method for the precise and complete correction of distortion on cineangiographic image: its effect on left ventricular measurements. *Catheterization & Cardiovascular Diagnosis*, 4(2):111–26, 1978.

[10] M. A. Borge and R. Marsan. Wedged hepatic venography with biplane digital subtraction imaging to facilitate accessing the portal venous system for tips [letter]. *Journal of Vascular & Interventional Radiology*, 9(5):853–4, 1998.

[11] N Alberto Borghese and P Cerveri. Calibrating a video camera pair with a rigid bar. *Pattern Recognition*, 33:81–95, 2000.

[12] E Bullitt, S Aylward, K Smith, S K Mukherji, M Jiroutek, and K Muller. Symbolic description of intracerebral vessels segmented from mra and evaluation by comparison with x-ray angiograms. *Medical Image Analysis*, 5:157–169, 2001.

[13] E Bullitt, A Liu, S Aylward, C Coffey, J Stone, S K Mukherji, K Muller, and S M Pizer. Registration of 3d vessels with 2d digital angiograms: clinical evaluation. *Academic Radiology*, 6:539–546, 1999.

[14] E. Bullitt, A. Liu, and S. M. Pizer. Three-dimensional reconstruction of curves from pairs of projection views in the presence of error. i. algorithms. *Medical Physics*, 24(11):1671–8, 1997.

[15] E. Bullitt, A. Liu, and S. M. Pizer. Three-dimensional reconstruction of curves from pairs of projection views in the presence of error. ii. analysis of error. *Medical Physics*, 24(11):1679–87, 1997.

[16] D. P. Chakraborty. Image intensifier distortion correction. *Medical Physics*, 14(2):249–52, 1987.

[17] S. Y. Chen and C. E. Metz. Improved determination of biplane imaging geometry from two projection images and its application to three-dimensional reconstruction of coronary arterial trees. *Medical Physics*, 24(5):633–54, 1997.

[18] M. A. Clifford, F. Banovac, E. Levy, and K. Cleary. Assessment of hepatic motion secondary to respiration for computer assisted interventions. *Computer Aided Surgery*, 7:291299, 2002.

[19] T.F. Coleman and Y. Li. On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67(2):189–224, 1994.

[20] T.F. Coleman and Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.

[21] G. Corral, K. Cleary, J. Tang, E. Levy, and N. Glossop. Orientation accuracy of a magnetic tracking device for image-guided interventions. In *CARS*, page 1367, 2003.

[22] S. C. Davies, A. L. Hill, R. B. Holmes, M. Halliwell, and P. C. Jackson. Ultrasound quantitation of respiratory organ motion in the upper abdomen. *Br J Radiol*, 67:1096–1102, 1994.

[23] Gerald M. Doherty and Lawrence W. Way. Liver & portal venous system. In Lawrence W. Way, editor, *Current Surgical Diagnosis & Treatment*. Appleton & Lange, 11th edition, 2003.

[24] P. J. Edwards, A. P. King, C. R. Jr. Maurer, D. A. deCunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. Jusczyzck, A. J. Strong, C. L. Chandler, and M. J. Gleeson. Design and evaluation of a system for microscope-assisted guided interventions (magi). *IEEE Transactions on Medical Imaging*, 19(11):1082–1093, 2000.

[25] R. Fahrig, M. Moreau, and D. W. Holdsworth. Three-dimensional computed tomographic reconstruction using a c-arm mounted xrii: Correction of image intensifier distortion. *Medical Physics*, 24(7):1097–1105, 1997.

[26] O.D. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.

[27] O.D. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings CVPR '86: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.86CH2290-5)*, pages 15–20, Washington, DC, USA, 1986. IEEE Comput. Soc. Press.

[28] Olivier Faugeras and M. Hebert. The representation, recognition, and locating of 3-d objects. *International Journal of Robotics Research*, 5(3):27–52, 1986.

[29] L. E. Fencil and C. E. Metz. Propagation and reduction of error in three-dimensional structure determined from biplane views of unknown orientation. *Medical Physics*, 17(6):951–61, 1990.

[30] David A. Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Pearson Education, Upper Saddle Rivier, New Jersey, 2003.

[31] A. M. Freedman, A. J. Sanyal, J. Tisnado, P. E. Cole, M. L. Shiffman, V. A. Luketic, P. P. Purdum, M. D. Darcy, and M. P. Posner. Complications of transjugular intrahepatic portosystemic shunt: a comprehensive review. *Radiographics*, 13(6):1185–210, 1993. Review.

[32] H. Fuchs, A. State, E. Pisano, W. Garrett, G. Hirota, M. Livingston, M. Whitton, and S. Pizer. Performing ultrasound-guided needle biopsies from within a head-mounted display. *Proceedings of Visualization in Biomedical Computing 1996*, pages 591–600, 1996.

[33] D. R. Ganger, J. B. Klapman, V. McDonald, T. A. Matalon, S. Kaur, H. Rosenblate, R. Kane, M. Saker, and D. M. Jensen. Transjugular intrahepatic portosystemic shunt (tips) for budd-chiari syndrome or portal vein thrombosis: review of indications and problems [see comments]. *American Journal of Gastroenterology*, 94(3):603–8, 1999.

[34] W. Garrett, H. Fuchs, M. Whitton, and A. State. Real-time incremental visualization of dynamic ultrasound volumes using parallel bsp trees. *Proceedings of IEEE Visualization 1996*, pages 235–240, 1996.

[35] E. Gronenschild. The accuracy and reproducibility of a global method to correct for geometric image distortion in the x-ray imaging chain. *Medical Physics*, 24(12):1875–88, 1997.

[36] E. Gronenschild. Correction for geometric image distortion in the x-ray imaging chain: local technique versus global technique. *Medical Physics*, 26(12):2602–16, 1999.

[37] P. Haaker, E. Klotz, R. Koppe, and R. Linde. Real-time distortion correction of digital x-ray ii/tv-systems: an application example for digital flashing tomosynthesis (dfts). *International Journal of Cardiac Imaging*, 6:39–45, 1991.

[38] Haralick and Shapiro. Appendix a - ellipse calculations. In *Computer and Robot Vision*, volume 1, pages 639–658. Addison-Wesley, 1992.

[39] A. M. Harper and J. D. Rosendale. The unos optn waiting list and donor registry: 1988-1996. *Clinical Transplants*, pages 69–90, 1996.

[40] T. Hartkens, D.L.G. Hill, A.D. Castellano-Smith, D.J. Hawkes, Jr. Maurer, C.R., A.J. Martin, W.A. Hall, H. Liu, and C.L. Truwit. Measurement and analysis of brain deformation during neurosurgery. *IEEE Transactions on Medical Imaging*, 22(1):82–92, 2003.

[41] Richard Hartley. Projective reconstruction from line correspondence. In *IEEE Conference on Computer Vision and Pattern Recognition*, 1994.

[42] Richard Hartley and Andrew Zisserman. *Multiple View Geometry*. Cambridge University Press, Cambridge, UK, 2000.

[43] A. J. Herline, J. D. Stefansic, J. P. Debelak, S. L. Hartmann, C. W. Pinson, R. L. Galloway, and W. C. Chapman. Image guided surgery: Preliminary feasibility studies of frameless stereotactic liver surgery. *Arch Surg*, 134:644–650, 1999.

[44] Berthold K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.

[45] M. Jacobs, M. Livingston, and A. State. Managing latency in complex augmented reality systems. *ACM SIGGRAPH Proceedings of the Symposium of Interactive 3D Graphics 1997, Providence, RI*, pages 49–54, 1997.

[46] F. A. Jolesz, L. Hsu, R. B. Schwartz, and et al. Mr-guided neurointerventions. In *Proceedings of American Society of Neuroradiology, 34th Annual Meeting*, page 218, Seattle, Washington, 1996.

[47] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, pages 35–45, 1960.

[48] R.E. Kalman. New methods in wiener filtering theory. In *1st Symposium of Engineering Application of Random Function Theory and Probability*, pages 270–388, New York, 1963. Wiley.

[49] R.E. Kalman. Pattern recognition properties of multilinear response function. ii. *Control and Cybernetics*, 9(1-2):21–31, 1980. (03) Stanford Univ., Stanford, CA, USA (39) CCYBAP.

[50] R.E. Kalman. Pattern recognition properties of multilinear response functions. i. *Control and Cybernetics*, 9(1-2):5–20, 1980. (03) Stanford Univ., Stanford, CA, USA (39) CCYBAP.

[51] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, pages 95–108, 1961.

[52] A.P. King, J.M. Blackall, G.P. Penney, and D.J. Hawkes. Tracking liver motion using 3-d ultrasound and a surface based statistical shape model. In L. Staib, editor, *Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 145–52, Los Alamitos, CA, USA Kauai, HI, USA, 2001. IEEE Comput. Soc.

[53] H. W. Korin, R. L. Ehman, S. J. Riederer, J. P. Felmlee, and R. C. Grimm. Respiratory kinematics of the upper abdominal organs: A quantitative study. *Magn Reson Med*, 23:172–178, 1992.

[54] H. D. Kubo and B. C. Hill. Respiration gated radiotherapy treatment: A technical study. *Phys Med Biol*, 41:8391, 1996.

[55] Jack B. Kuipers. *Quaternions and Rotation Sequences*. Princeton University Press, Princeton, NJ, 1999.

[56] E. Levy, M. Kalis, M. Vo, D. Lindisch, and K. Cleary. Feasibility of simultaneous respiratory function monitoring and determination of respiratory-related intrahepatic vessel excursion using the lifeshirt^{TM}system. In *CARS*, 2004.

[57] A Liu and E Bullitt. 3d/2d registration using tubular anatomical structures as a basis. *submitted IEEE-TMI 6/99*, 1999.

[58] A. Liu, E. Bullitt, and S.M. Pizer. 3d/2d registration via skeletal near projective invariance in tubular objects. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98. First International Conference. Proceedings*, pages 952–63, 1998. (03) Center for Inf.-Enhanced Med., Nat. Univ. of Singapore, Singapore.

[59] Alan Ve-Ming Liu. *3D/2D Registration and Reconstruction in Image-Guided Surgery.* PhD thesis, University of North Carolina at Chapel Hill, 1998.

[60] R. R. Liu, S. Rudin, and D. R. Bednarek. Super-global distortion correction for a rotational c-arm x-ray image intensifier. *Medical Physics*, 26(9):1802–10, 1999.

[61] H.C. Longuet-Higgins. A computer algorithm for recon- structing a scene from two projections. *Nature*, pages 133–135, 1981.

[62] C. R. Jr. Maurer, F. Sauer, B. Hu, B. Bascle, B. Geiger, F. Wenzel, F. Recchi, T. Rohlfing, C. M. Brown, R. S. Bakos, R. J. Maciunas, and A. Bani-Hashemi. Augmented reality visualization of brain structures with stereo and kinetic depth cues: System description and initial evaluation with head phantom. presented at medical imaging 2001: Visualization, display, and image-guided procedures. *Proceedings of SPIE*, 4319:445–456, 2001.

[63] Robert A. McLaughlin, John Hipwell, Graeme P. Penney, Kawal Rhode, Albert Chung, J. Alison Noble, and D. J. Hawkes. Intensity-based registration versus feature-based registration for neurointerventions. *Proceedings of Medical Image Understanding and Analysis 2001, University of Birmingham, UK*, 2001.

[64] C. E. Metz and L. E. Fencil. Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: theory. *Medical Physics*, 16(1):45–51, 1989.

[65] D.L. Miller, S. Balter, P.E. Cole, and H.T. et. al. Lu. Radiation doses in interventional radiology procedures: The rad-ir study. part i: Overall measures of dose. *J Vasc Interv Radiol*, 14:711–727, 2003.

[66] D.L. Miller, S. Balter, P.E. Cole, and H.T. et. al. Lu. Radiation doses in interventional radiology procedures: The rad-ir study. part ii: Skin dose. *J Vasc Interv Radiol*, 14:977–990, 2003.

[67] M. Mitschke, A. Bani-Hashemi, and N. Navab. Interventions under video-augmented x-ray guidance: Application to needle placement. In *MICCAI*, pages 858–868, 2000.

[68] Frank H. Netter. *Atlas of Human Anatomy.* Novartis, East Hanover, New Jersey, 2nd edition, 1997.

[69] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[70] Nikos Paragios. *Geodesic Active Regions and Level Set Methods: Contributions and Applications in Artificial Vision.* PhD thesis, INRIA Sophia-Antipolis, 2000.

[71] Nikos Paragios and Rachid Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3):266–280, 2000.

[72] S. H. Parker, A. T. Stavros, and M. A. Dennis. Needle biopsy techniques. *Radiologic Clinics of North America*, 33(6):1171–86, 1995. Review.

[73] G. P. Penney, J. A. Little, D. L. G. Hill, and D. J. Hawkes. 2d-3d image registration for use in image guided interventions. *Proceedings of Medical Image Understanding and Analysis, University of Oxford, UK, July 1997*, 1997.

[74] Graeme P. Penney, Jrgen Weese, John A. Little, Paul Desmedt, Derek L. G. Hill, and David J. Hawkes. A comparison of similarity measures for use in 2-d–3-d medical image registration. *IEEE Transactions on Medical Imaging*, 17(4):586–95, 1998.

[75] E. D. Pisano, L. L. Fajardo, D. J. Caudry, N. Sneige, W. J. Frable, W. A. Berg, I. Tocino, S. J. Schnitt, J. L. Connolly, C. A. Gatsonis, and B. J. McNeil. Fine-needle aspiration biopsy of nonpalpable breast lesions in a multicenter clinical trial: results from the radiologic diagnostic oncology group v. *Radiology*, 219(3):785–92, 2001.

[76] M Pollefeys, R Koch, and L Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *International Journal of Computer Vision*, 32(1):7–25, 1999.

[77] M Pollefeys, R Koch, M Vergauwen, and L Van Gool. Auomated reconstruction of 3d scenes from sequences of images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(4):251–267, 2000.

[78] C. R. Rees, R. L. Niblett, S. P. Lee, N. G. Diamond, and J. S. Crippin. Use of carbon dioxide as a contrast medium for transjugular intrahepatic portosystemic shunt procedures. *Journal of Vascular & Interventional Radiology*, 5(2):383–6, 1994.

[79] S. C. Rose, D. H. Pretorius, T. B. Kinney, T. R. Nelson, K. Valji, H. R. D'Agostino, N. M. Forsythe, A. C. Roberts, and M. L. Manco-Johnson. Three-dimensional sonographic guidance for transvenous intrahepatic invasive procedures: feasibility of a new technique. *Journal of Vascular & Interventional Radiology*, 10(2 Pt 1):189–98, 1999.

[80] M. Rosenthal and H. Fuchs. X-ray phantom for simultaneous calibration of multiple views. *American Association of Physicists in Medicine (AAPM) Annual Meeting*, 2004.

[81] M. Rosenthal, A. State, J. Lee, G. Hirota, J. Ackerman, K. Keller, E. Pisano, M. Jiroutek, K. Muller, and H. Fuchs. Augmented reality guidance for needle biopsies: A randomized, controlled trial in phantoms. *Proceedings of MICCAI 2001, eds. W. Niessen and M. Viergever. Lecture Notes in Computer Science*, 2208:240–248, 2001.

[82] M. Rosenthal, S. Weeks, S. Aylward, E. Bullitt, and H. Fuchs. Intraoperative tracking of anatomical structures using fluoroscopy and a vascular balloon catheter. *Proceedings of MICCAI 2001, eds. W. Niessen and M. Viergever. Lecture Notes in Computer Science*, 2208:1253–4, 2001.

[83] Michael Rosenthal, Andrei State, Joohi Lee, Gentaro Hirota, Jeremy Ackerman, Kurtis Keller, Etta D. Pisano, Michael Jiroutek, Keith Muller, and Henry Fuchs. Augmented reality guidance for needle biopsies: An initial randomized, controlled trial in phantoms. *Medical Image Analysis*, 6(3):313–320, 2002. TY - JOUR.

[84] S. Rudin, D. R. Bednarek, and R. Wong. Accurate characterization of image intensifier distortion. *Medical Physics*, 18(6):1145–51, 1991.

[85] F. Sauer, A. Khamene, and S. Vogt. An augmented reality navigation system with a single-camera tracker: System design and needle biopsy phantom trial. *Lecture Notes in Computer Science*, 2489:116–124, 2002.

[86] R. R. Saxon and F. S. Keller. Technical aspects of accessing the portal vein during the tips procedure. *Journal of Vascular & Interventional Radiology*, 8(5):733–44, 1997. Review.

[87] J. F. Schenck, F. A. Jolesz, P. B. Roemer, and et al. Superconducting open configuration mri system for image-guided therapy. *Radiology*, 195:805–814, 1995.

[88] S. Shimizu, H. Shirato, H. Aoyama, S. Hashimoto, T. Nishioka, A. Yamazaki, K. Kagei, and K. Miyasaka. High speed magnetic resonance imaging for four-dimensional treatment planning of conformal radiotherapy of moving body tumors. *Int J Radiat Oncol Biol Phys*, 48:4714, 2000.

[89] S. Shimizu, H. Shirato, B. Xo, K. Kagei, T. Nishioka, S. Hashimoto, K. Tsuchiya, H. Aoyama, and K. Miyasaka. Three-dimensional movement of a liver tumor detected by high-speed magnetic resonance imaging. *Radiother Onco*, 50:367–370, 1999.

[90] S. G. Silverman, B. D. Collick, M. R. Figueira, R. Khorasani, D. F. Adams, R. W. Newman, G. P. Topulos, and F. A. Jolesz. Interactive mr-guided biopsy in an open-configuration mr imaging system. *Radiology*, 197:173–181, 1995.

[91] M. Söderman, C. Picard, and K. Ericson. An algorithm for correction of distortion in stereotaxic digital subtraction angiography. *Neuroradiology*, 40:277–282, 1998.

[92] A. State, J. Ackerman, G. Hirota, J. Lee, and H. Fuchs. Dynamic virtual convergence for video see-through head-mounted displays: Maintaining maximum stereo overlap throughout a close-range work space. *Proc. International Symposium on Augmented Reality 2001, New York, NY*, pages 137–146, 2001.

[93] A. State, D. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, and H. Fuchs. Case study: Observing a volume-rendered fetus within a pregnant patient. *Proceedings of IEEE Visualization 1994, Los Alamitos, CA*, pages 364–368, 1994.

[94] A. State, G. Hirota, D. Chen, W. Garrett, and M. Livingston. Superior augmented-reality registration by integrating landmark tracking and magnetic tracking. *ACM SIGGRAPH Proceedings of the Symposium of Interactive 3D Graphics 1997, Providence, RI*, pages 429–438, 1996.

[95] A. State, K. Keller, M. Rosenthal, H. Yang, J. Ackerman, , and H. Fuchs. Stereo imagery from the unc augmented reality system for breast biopsy guidance. In *Medicine Meets Virtual Reality (MMVR)*, Newport Beach, CA, 2003.

[96] G. Stetten and V. Chib. Overlaying ultrasound images on direct vision. *Journal of Ultrasound in Medicine*, 20(1):235–240, 2001.

[97] I. Sutherland. A head-mounted three-dimensional display. *AFIPS Conference Proceedings*, 33-I:757–764, 1968.

[98] Ivan Sutherland. Sketchpad: A man-machine graphical communications system. Technical Report Technical Report 296, MIT Lincoln Laboratories, 1963.

[99] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. Bundle adjustment - a modern synthesis. In Bill Triggs, A. Zisserman, and R. Szeliski, editors, *Vision Algorithms: Theory & Practice*, volume LNCS 1883. Springer-Verlag, 2000.

[100] W Triggs. The geometry of projective reconstruction i: Matching constraints and the joint image. In *International Conference on Computer Vision*, pages 338–343, 1995.

[101] E. Trucco and A. Verri. Camera calibration. In *Introductory Techniques for 3-D Computer Vision*, pages 123–138. Prentice-Hall, Inc., Upper Saddle River, NJ, 1998.

[102] R.Y. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. *Proceedings CVPR '86: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.86CH2290-5)*, pages 364–74, 1986. (03) IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, USA.

[103] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–44, 1987. (03) IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, USA (39) IJRAE4.

[104] Susan Weeks1 Vijay Venkatraman2, Mark H. Van Horn1 and Elizabeth Bullitt. Liver motion due to needle pressure, cardiac and respiratory motion during the tips procedure. *Lecture Notes in Computer Science, Proceedings of MICCAI 2004*, 3217:66–72, 2004.

[105] William A. Wallace and Frank Johnson. Detection and correction of geometrical distortion in x-ray fluoroscopic images. *Biomechanics*, 14:123–125, 1981.

[106] G. Welch and G. Bishop. An introduction to the kalman filter. Technical Report TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, 2001.

[107] B. J. Wood, D. Lindisch, S. Ranjan, N. Glossop, and K. Cleary. Electromagnetically tracked guidewires for interventional procedures. In *CARS*, 2004.

[108] Jae-Woong Yi, Tae Seok Yang, and Jun-Ho Oh. Estimation of depth and 3d motion parameters of moving objects with multiple stereo images by using kalman filter. *Proceedings of IECON '95 - 21st Annual Conference on IEEE Industrial Electronics*, 2:1225–30, 1995. (03) Dept. of Mech. Eng., Korea Adv. Inst. of Sci. & Technol., Taejon, South Korea.

Appendix A

PARTIAL DERIVATIVES OF CALIBRATION COST FUNCTIONS

This section contains the partial derivative equations used in the calibration bundle adjustment technique described in Chapter 4.

The pinhole camera projection model used in this work is described by the following equations:

$$\begin{bmatrix} x_{ij}^{proj} \\ y_{ij}^{proj} \\ z_{ij}^{proj} \end{bmatrix} = \begin{bmatrix} f_{x_i} & sk_i & ocx_i \\ 0 & f_{y_i} & ocy_i \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_i^{11} & r_i^{12} & r_i^{13} & T_{x_i} \\ r_i^{21} & r_i^{22} & r_i^{23} & T_{y_i} \\ r_i^{31} & r_i^{32} & r_i^{33} & T_{z_i} \end{bmatrix} \begin{bmatrix} x_j^{mod} \\ y_j^{mod} \\ z_j^{mod} \\ 1 \end{bmatrix} \tag{A.1}$$

$$X_{ij}^{proj} = L_i \left[R_i \mid T_i \right] X_j^{mod} \tag{A.2}$$

$$X_{ij}^{proj} = L_i M_i X_j^{mod} \tag{A.3}$$

$$X_{ij}^{proj} = P_i X_j^{mod} \tag{A.4}$$

where x_{ij}^{proj}, y_{ij}^{proj}, and z_{ij}^{proj} are unnormalized homogeneous coordinates. The normalized coordinates $\hat{x}_{ij}^{proj} = \frac{x_{ij}^{proj}}{z_{ij}^{proj}}$ and $\hat{y}_{ij}^{proj} = \frac{y_{ij}^{proj}}{z_{ij}^{proj}}$ are used to calculate within-image residual errors.

When one considers a system with two cameras, it becomes desirable to describe the relationship between them and the world coordinates explicitly via the extrinsic transformations, as follows:

$$X^A = \begin{bmatrix} R^{W \to A} & T^{W \to A} \end{bmatrix} X^W = M^{W \to A} X^W \tag{A.5}$$

$$X^B = \begin{bmatrix} R^{W \to B} & T^{W \to B} \end{bmatrix} X^W \tag{A.6}$$

$$= \begin{bmatrix} R^{A \to B} & T^{A \to B} \end{bmatrix} \begin{bmatrix} R^{W \to A} & T^{W \to A} \end{bmatrix} X^W \tag{A.7}$$

$$= M^{A \to B} M^{W \to A} X^W \tag{A.8}$$

The transformation matrices are expanded to 4x4 homogeneous coordinate systems as needed to allow for chain multiplication (i.e., $M^{W \to A}$ would have the vector [0 0 0 1] appended to its lower boundary when used in the multiplication chain $M^{A \to B} M^{W \to A} X^W$). The projection

equations for the two-camera system can thus be described as:

$$X_{ij}^{A} = L_i M_i^{W \to A} X_j^{mod} = P_i^A X_j^{mod} \quad \text{and} \tag{A.9}$$

$$X_{ij}^{A} = L_i M^{A \to B} M_i^{W \to A} X_j^{mod} = P_i^B X_j^{mod} \tag{A.10}$$

The least-squares function to be optimized, $Cost(\mathbf{x})$, is described below. This cost function has three components: $Err(\mathbf{x})$, the sum of the geometric error distances between each image point and the projection of the corresponding model point; $Var(\mathbf{x})$, the weighted sum of the variances of the parameters that are being globalized; and $Constr(\mathbf{x})$, a cost element intended to restrict the correction of the model points. In these equations, m represents the number of frame pairs, n represents the number of model points, $P_i^{A|B}$ represents the projection matrix for camera A or B, X_j^{mod} represents the three-dimensional coordinates of the j^{th} model point, and $x_j^{A|B}$ represents the 2D image coordinate of the j^{th} point on camera A or B. In addition, Γ represents the set of parameters to be globalized, as described above, p_{li} represents the l^{th} such parameter in the i^{th} frame, and μ_l is the mean of the l^{th} parameters across all frames. $\mathbf{x}$ is a linear vector of all of the parameters of the system.

$$Cost(\mathbf{x}) = Err(\mathbf{x}) + Var(\mathbf{x}) + Constr(\mathbf{x}) \tag{A.11}$$

$$Err(\mathbf{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(P_i^A X_j^{mod} - x_j^A\right)^2 + \left(P_i^B X_j^{mod} - x_j^B\right)^2 \tag{A.12}$$

$$Var(\mathbf{x}) = w_l \sum_{l \in \Gamma} \frac{\sum_{i=1}^{m} \left(p_{li} - \mu_l\right)^2}{m-1} \tag{A.13}$$

$$Constr(\mathbf{x}) = m \sum_{j=1}^{n} \left(\kappa_{x_j}^2 + \kappa_{y_j}^2 + \kappa_{z_j}^2\right) \tag{A.14}$$

First, let us consider the Jacobian (matrix of first derivatives) of the $Err(\mathbf{x})$ component of this cost function. For the purpose of simplification, consider this error function to be a sum

of residual errors in each view, as follows:

$$Err(\mathbf{x}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left({\rho_{ij}^{A}}^2 + {\rho_{ij}^{B}}^2\right) \quad \text{where} \tag{A.15}$$

$$\rho_{ij}^{A} = \sqrt{\left(\frac{P_i^{A1}X_j^{mod}}{P_i^{A3}X_j^{mod}} - x_j^A\right)^2 + \left(\frac{P_i^{A2}X_j^{mod}}{P_i^{A3}X_j^{mod}} - y_j^A\right)^2} \quad \text{and} \tag{A.16}$$

$$\rho_{ij}^{B} = \sqrt{\left(\frac{P_i^{B1}X_j^{mod}}{P_i^{B3}X_j^{mod}} - x_j^B\right)^2 + \left(\frac{P_i^{B2}X_j^{mod}}{P_i^{B3}X_j^{mod}} - y_j^B\right)^2} \tag{A.17}$$

are the reprojection errors (in pixels) associated with point i of frame j in camera A and B respectively. P_i^{An} represents the n^{th} row of the projection matrix P_i^A.

The numerical optimizer used in this work accepts a list of the *unsquared* cost elements to be optimized, so each ρ_{ij} must be considered separately. The elements of the jacobian matrix for each of these elements is as follows:

$$\frac{\partial \rho_{ij}^{A}}{\partial f_{x_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^{A} \right] \frac{M_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.18}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial f_{y_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^{A} \right] \frac{M_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.19}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial ocx_k} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^{A} \right] \quad \text{if} \quad i = k \end{cases} \tag{A.20}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial ocy_k} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^{A} \right] \quad \text{if} \quad i = k \end{cases} \tag{A.21}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial sk_k} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^{A} \right] \frac{M_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.22}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial T_{x_k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^{A} \right] \frac{f_{x_i}}{P_i^{A3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.23}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial T_{y_k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^{A} \right) \frac{sk_i}{P_i^{A3} X_j^{mod}} + \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^{A} \right) \frac{f_{x_i}}{P_i^{A3} X_j^{mod}} \right] \quad \text{if} \quad i = k \end{cases} \tag{A.24}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial T_{z_k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^{A} \right) \left(\frac{ocx_i}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^{A} \right) \left(\frac{ocy_i}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.25}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial \alpha_{k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^A \right) \left(\frac{L_i^{A1} \frac{\partial R_i}{\partial \alpha_i} M_i^A X_j^{mod}}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod} L_i^{A3} \frac{\partial R_i}{\partial \alpha_i} M_i^A X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^A \right) \left(\frac{L_i^{A2} \frac{\partial R_i}{\partial \alpha_i} M_i^A X_j^{mod}}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod} L_i^{A3} \frac{\partial R_i}{\partial \alpha_i} M_i^A X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.26}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial \beta_{k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^A \right) \left(\frac{L_i^{A1} \frac{\partial R_i}{\partial \beta_i} M_i^A X_j^{mod}}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod} L_i^{A3} \frac{\partial R_i}{\partial \beta_i} M_i^A X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^A \right) \left(\frac{L_i^{A2} \frac{\partial R_i}{\partial \beta_i} M_i^A X_j^{mod}}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod} L_i^{A3} \frac{\partial R_i}{\partial \beta_i} M_i^A X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.27}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial \gamma_{k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^A \right) \left(\frac{L_i^{A1} \frac{\partial R_i}{\partial \gamma_i} M_i^A X_j^{mod}}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod} L_i^{A3} \frac{\partial R_i}{\partial \gamma_i} M_i^A X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^A \right) \left(\frac{L_i^{A2} \frac{\partial R_i}{\partial \gamma_i} M_i^A X_j^{mod}}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod} L_i^{A3} \frac{\partial R_i}{\partial \gamma_i} M_i^A X_j^{mod}}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.28}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial \kappa_{x_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^A \right) \left(\frac{L_i^{A1} R_i [1 \ 0 \ 0]^T}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod} L_i^{A3} R_i [1 \ 0 \ 0]^T}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^A \right) \left(\frac{L_i^{A2} R_i [1 \ 0 \ 0]^T}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod} L_i^{A3} R_i [1 \ 0 \ 0]^T}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.29}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial \kappa_{y_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^A \right) \left(\frac{L_i^{A1} R_i [0 \ 1 \ 0]^T}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod} L_i^{A3} R_i [0 \ 1 \ 0]^T}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^A \right) \left(\frac{L_i^{A2} R_i [0 \ 1 \ 0]^T}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod} L_i^{A3} R_i [0 \ 1 \ 0]^T}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.30}$$

$$\frac{\partial \rho_{ij}^{A}}{\partial \kappa_{z_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{A}} \left[\left(\frac{P_i^{A1} X_j^{mod}}{P_i^{A3} X_j^{mod}} - x_j^A \right) \left(\frac{L_i^{A1} R_i [0 \ 0 \ 1]^T}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A1} X_j^{mod} L_i^{A3} R_i [0 \ 0 \ 1]^T}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{A2} X_j^{mod}}{P_i^{A3} X_j^{mod}} - y_j^A \right) \left(\frac{L_i^{A2} R_i [0 \ 0 \ 1]^T}{P_i^{A3} X_j^{mod}} - \frac{P_i^{A2} X_j^{mod} L_i^{A3} R_i [0 \ 0 \ 1]^T}{\left(P_i^{A3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.31}$$

For camera B (lateral view), the equations of the jacobian elements are as follows:

$$\frac{\partial \rho_{ij}^{B}}{\partial f_{x_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right] \frac{M_i^{B1} M_i^{A} X_j^{mod}}{P_i^{B3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.32}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial f_{y_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right] \frac{M_i^{B2} M_i^{A} X_j^{mod}}{P_i^{B3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.33}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial ocx_k} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right] \quad \text{if} \quad i = k \end{cases} \tag{A.34}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial ocy_k} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right] \quad \text{if} \quad i = k \end{cases} \tag{A.35}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial sk_k} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right] \frac{M_i^{B2} M_i^{A} X_j^{mod}}{P_i^{B3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.36}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial T_{x_k}^{B}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right] \frac{f_{x_i}}{P_i^{B3} X_j^{mod}} \quad \text{if} \quad i = k \end{cases} \tag{A.37}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial T_{y_k}^{B}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right) \frac{sk_i}{P_i^{B3} X_j^{mod}} + \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right) \frac{f_{x_i}}{P_i^{B3} X_j^{mod}} \right] \quad \text{if} \quad i = k \end{cases} \tag{A.38}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial T_{z_k}^{B}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right) \left(\frac{ocx_i}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right) \left(\frac{ocy_i}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.39}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial T_{x_k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right) \left(\frac{L_i^{B1} R_i^{A \to B.1}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} R_i^{A \to B.1}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right) \left(\frac{L_i^{B2} R_i^{A \to B.1}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} R_i^{A \to B.1}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.40}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial T_{y_k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right) \left(\frac{L_i^{B1} R_i^{A \to B.2}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} R_i^{A \to B.2}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right) \left(\frac{L_i^{B2} R_i^{A \to B.2}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} R_i^{A \to B.2}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.41}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial T_{z_k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^{B} \right) \left(\frac{L_i^{B1} R_i^{A \to B.3}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} R_i^{A \to B.3}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^{B} \right) \left(\frac{L_i^{B2} R_i^{A \to B.3}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} R_i^{A \to B.3}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.42}$$

(A.43)

$$\frac{\partial \rho_{ij}^{B}}{\partial \alpha_{k}^{B}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^B \right) \left(\frac{L_i^{B1} \frac{\partial R_i^{A \to B}}{\partial \alpha_i^B} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} \frac{\partial R_i^{A \to B}}{\partial \alpha_i^B} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^B \right) \left(\frac{L_i^{B2} \frac{\partial R_i^{A \to B}}{\partial \alpha_i^B} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} \frac{\partial R_i^{A \to B}}{\partial \alpha_i^B} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.44}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial \beta_{k}^{B}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^B \right) \left(\frac{L_i^{B1} \frac{\partial R_i^{A \to B}}{\partial \beta_i^B} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} \frac{\partial R_i^{A \to B}}{\partial \beta_i^B} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^B \right) \left(\frac{L_i^{B2} \frac{\partial R_i^{A \to B}}{\partial \beta_i^B} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} \frac{\partial R_i^{A \to B}}{\partial \beta_i^B} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.45}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial \gamma_{k}^{B}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^B \right) \left(\frac{L_i^{B1} \frac{\partial R_i^{A \to B}}{\partial \gamma_i^B} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} \frac{\partial R_i^{A \to B}}{\partial \gamma_i^B} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^B \right) \left(\frac{L_i^{B2} \frac{\partial R_i^{A \to B}}{\partial \gamma_i^B} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} \frac{\partial R_i^{A \to B}}{\partial \gamma_i^B} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.46}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial \alpha_{k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^B \right) \left(\frac{L_i^{B1} \frac{\partial R_i^{W \to A}}{\partial \alpha_i^A} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} \frac{\partial R_i^{W \to A}}{\partial \alpha_i^A} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^B \right) \left(\frac{L_i^{B2} \frac{\partial R_i^{W \to A}}{\partial \alpha_i^A} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} \frac{\partial R_i^{W \to A}}{\partial \alpha_i^A} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.47}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial \beta_{k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^B \right) \left(\frac{L_i^{B1} \frac{\partial R_i^{W \to A}}{\partial \beta_i^A} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} \frac{\partial R_i^{W \to A}}{\partial \beta_i^A} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^B \right) \left(\frac{L_i^{B2} \frac{\partial R_i^{W \to A}}{\partial \beta_i^A} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} \frac{\partial R_i^{W \to A}}{\partial \beta_i^A} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.48}$$

$$\frac{\partial \rho_{ij}^{B}}{\partial \gamma_{k}^{A}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^{B}} \left[\left(\frac{P_i^{B1} X_j^{mod}}{P_i^{B3} X_j^{mod}} - x_j^B \right) \left(\frac{L_i^{B1} \frac{\partial R_i^{W \to A}}{\partial \gamma_i^A} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B1} X_j^{mod} L_i^{B3} \frac{\partial R_i^{W \to A}}{\partial \gamma_i^A} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) + \right. \\ \left. \left(\frac{P_i^{B2} X_j^{mod}}{P_i^{B3} X_j^{mod}} - y_j^B \right) \left(\frac{L_i^{B2} \frac{\partial R_i^{W \to A}}{\partial \gamma_i^A} X_j^{mod}}{P_i^{B3} X_j^{mod}} - \frac{P_i^{B2} X_j^{mod} L_i^{B3} \frac{\partial R_i^{W \to A}}{\partial \gamma_i^A} X_j^{mod}}{\left(P_i^{B3} X_j^{mod} \right)^2} \right) \right] \quad \text{if} \quad i = k \end{cases} \tag{A.49}$$

$$\frac{\partial \rho_{ij}^B}{\partial \kappa_{x_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^B}\left[\left(\frac{P_i^{B1}X_j^{mod}}{P_i^{B3}X_j^{mod}} - x_j^B\right)\left(\frac{L_i^{B1}R_i^{W\to B}[1\ 0\ 0]^T}{P_i^{B3}X_j^{mod}} - \frac{P_i^{B1}X_j^{mod}L_i^{B3}R_i^{W\to B}[1\ 0\ 0]^T}{\left(P_i^{B3}X_j^{mod}\right)^2}\right) + \right. \\ \left.\left(\frac{P_i^{B2}X_j^{mod}}{P_i^{B3}X_j^{mod}} - y_j^B\right)\left(\frac{L_i^{B2}R_i^{W\to B}[1\ 0\ 0]^T}{P_i^{B3}X_j^{mod}} - \frac{P_i^{B2}X_j^{mod}L_i^{B3}R_i^{W\to B}[1\ 0\ 0]^T}{\left(P_i^{B3}X_j^{mod}\right)^2}\right)\right] \quad \text{if} \quad i = k \end{cases} \tag{A.50}$$

$$\frac{\partial \rho_{ij}^B}{\partial \kappa_{y_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^B}\left[\left(\frac{P_i^{B1}X_j^{mod}}{P_i^{B3}X_j^{mod}} - x_j^B\right)\left(\frac{L_i^{B1}R_i^{W\to B}[0\ 1\ 0]^T}{P_i^{B3}X_j^{mod}} - \frac{P_i^{B1}X_j^{mod}L_i^{B3}R_i^{W\to B}[0\ 1\ 0]^T}{\left(P_i^{B3}X_j^{mod}\right)^2}\right) + \right. \\ \left.\left(\frac{P_i^{B2}X_j^{mod}}{P_i^{B3}X_j^{mod}} - y_j^B\right)\left(\frac{L_i^{B2}R_i^{W\to B}[0\ 1\ 0]^T}{P_i^{B3}X_j^{mod}} - \frac{P_i^{B2}X_j^{mod}L_i^{B3}R_i^{W\to B}[0\ 1\ 0]^T}{\left(P_i^{B3}X_j^{mod}\right)^2}\right)\right] \quad \text{if} \quad i = k \end{cases} \tag{A.51}$$

$$\frac{\partial \rho_{ij}^B}{\partial \kappa_{z_k}} = \begin{cases} 0 \quad \text{if} \quad i \neq k \\ \frac{1}{\rho_{ij}^B}\left[\left(\frac{P_i^{B1}X_j^{mod}}{P_i^{B3}X_j^{mod}} - x_j^B\right)\left(\frac{L_i^{B1}R_i^{W\to B}[0\ 0\ 1]^T}{P_i^{B3}X_j^{mod}} - \frac{P_i^{B1}X_j^{mod}L_i^{B3}R_i^{W\to B}[0\ 0\ 1]^T}{\left(P_i^{B3}X_j^{mod}\right)^2}\right) + \right. \\ \left.\left(\frac{P_i^{B2}X_j^{mod}}{P_i^{B3}X_j^{mod}} - y_j^B\right)\left(\frac{L_i^{B2}R_i^{W\to B}[0\ 0\ 1]^T}{P_i^{B3}X_j^{mod}} - \frac{P_i^{B2}X_j^{mod}L_i^{B3}R_i^{W\to B}[0\ 0\ 1]^T}{\left(P_i^{B3}X_j^{mod}\right)^2}\right)\right] \quad \text{if} \quad i = k \end{cases} \tag{A.52}$$

This work uses the Euler yaw-pitch-roll angle format for rotations, yielding composite rotations of the form:

$$R_i = \begin{bmatrix} \cos(\alpha_i)\cos(\gamma_i) & \cos(\alpha_i)\sin(\gamma_i) & \sin(\alpha_i) \\ -\sin(\alpha_i)\sin(\beta_i)\cos(\gamma_i) - \cos(\beta_i)\sin(\gamma_i) & -\sin(\alpha_i)\sin(\beta_i)\sin(\gamma_i) + \cos(\beta_i)\cos(\gamma_i) & \cos(\alpha_i)\sin(\beta_i) \\ -\sin(\alpha_i)\cos(\beta_i)\cos(\gamma_i) + \sin(\beta_i)\sin(\gamma_i) & -\sin(\alpha_i)\cos(\beta_i)\sin(\gamma_i) - \sin(\beta_i)\cos(\gamma_i) & \cos(\alpha_i)\cos(\beta_i) \end{bmatrix} \tag{A.53}$$

The partial derivatives of these rotations are:

$$\frac{\partial R_i}{\partial \alpha_i} = \begin{bmatrix} -\sin(\alpha_i)\cos(\gamma_i) & -\sin(\alpha_i)\sin(\gamma_i) & \cos(\alpha_i) \\ -\cos(\alpha_i)\sin(\beta_i)\cos(\gamma_i) & -\cos(\alpha_i)\sin(\beta_i)\sin(\gamma_i) & -\sin(\alpha_i)\sin(\beta_i) \\ -\cos(\alpha_i)\cos(\beta_i)\cos(\gamma_i) & -\cos(\alpha_i)\cos(\beta_i)\sin(\gamma_i) & -\sin(\alpha_i)\cos(\beta_i) \end{bmatrix}$$

$$\frac{\partial R_i}{\partial \beta_i} = \begin{bmatrix} 0 & 0 & 0 \\ -\sin(\alpha_i)\cos(\beta_i)\cos(\gamma_i) + \sin(\beta_i)\sin(\gamma_i) & -\sin(\alpha_i)\cos(\beta_i)\sin(\gamma_i) - \sin(\beta_i)\cos(\gamma_i) & \cos(\alpha_i) \\ \sin(\alpha_i)\sin(\beta_i)\cos(\gamma_i) + \cos(\beta_i)\sin(\gamma_i) & \sin(\alpha_i)\sin(\beta_i)\sin(\gamma_i) - \cos(\beta_i)\cos(\gamma_i) & -\cos(\alpha_i) \end{bmatrix}$$

$$\frac{\partial R_i}{\partial \gamma_i} = \begin{bmatrix} -\cos(\alpha_i)\sin(\gamma_i) & \cos(\alpha_i)\cos(\gamma_i) & 0 \\ \sin(\alpha_i)\sin(\beta_i)\sin(\gamma_i) - \cos(\beta_i)\cos(\gamma_i) & -\sin(\alpha_i)\sin(\beta_i)\cos(\gamma_i) - \cos(\beta_i)\sin(\gamma_i) & 0 \\ \sin(\alpha_i)\cos(\beta_i)\sin(\gamma_i) + \sin(\beta_i)\cos(\gamma_i) & -\sin(\alpha_i)\cos(\beta_i)\cos(\gamma_i) + \sin(\beta_i)\sin(\gamma_i) & 0 \end{bmatrix}$$

Distortion correction adds an additional set of entries to the jacobian matrix. This work uses Gronenschild's fifth-order polynomial method for distortion correction, which is described by the abbreviated equations:

$$\hat{x} = k^x_{20}x^2 + k^x_{11}xy + k^x_{02}y^2 + k^x_{30}x^3 + k^x_{21}x^2y + \cdots + k^x_{14}xy^4 + k^x_{05}y^5 \quad \text{(A.57)}$$

$$\hat{y} = k^y_{20}x^2 + k^y_{11}xy + k^y_{02}y^2 + k^y_{30}x^3 + k^y_{21}x^2y + \cdots + k^y_{14}xy^4 + k^y_{05}y^5 \quad \text{(A.58)}$$

When distortion correction is enabled, the corrected image coordinates $\hat{x}$ and $\hat{y}$ replace the image coordinates x and y throughout the cost equations and jacobian matrices. The effects of the coefficients $k^x_{\cdot\cdot}$ and $k^y_{\cdot\cdot}$ on the residual error costs are described by the equations below:

$$\frac{\partial \rho_{ij}}{\partial k^x_{20}} = -\frac{1}{\rho_{ij}}\left[\frac{P^1_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{x}_j\right] x_j^2 \quad \text{(A.59)}$$

$$\frac{\partial \rho_{ij}}{\partial k^x_{11}} = -\frac{1}{\rho_{ij}}\left[\frac{P^1_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{x}_j\right] x_j y_j \quad \text{(A.60)}$$

$$\frac{\partial \rho_{ij}}{\partial k^x_{02}} = -\frac{1}{\rho_{ij}}\left[\frac{P^1_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{x}_j\right] y_j^2 \quad \text{(A.61)}$$

$$\vdots \quad \text{(A.62)}$$

$$\frac{\partial \rho_{ij}}{\partial k^x_{14}} = -\frac{1}{\rho_{ij}}\left[\frac{P^1_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{x}_j\right] x_j y_j^4 \quad \text{(A.63)}$$

$$\frac{\partial \rho_{ij}}{\partial k^x_{05}} = -\frac{1}{\rho_{ij}}\left[\frac{P^1_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{x}_j\right] y_j^5 \quad \text{(A.64)}$$

and (A.65)

$$\frac{\partial \rho_{ij}}{\partial k^y_{20}} = -\frac{1}{\rho_{ij}}\left[\frac{P^2_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{y}_j\right] x_j^2 \quad \text{(A.66)}$$

$$\frac{\partial \rho_{ij}}{\partial k^y_{11}} = -\frac{1}{\rho_{ij}}\left[\frac{P^2_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{y}_j\right] x_j y_j \quad \text{(A.67)}$$

$$\frac{\partial \rho_{ij}}{\partial k^y_{02}} = -\frac{1}{\rho_{ij}}\left[\frac{P^2_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{y}_j\right] y_j^2 \quad \text{(A.68)}$$

$$\vdots \quad \text{(A.69)}$$

$$\frac{\partial \rho_{ij}}{\partial k^y_{14}} = -\frac{1}{\rho_{ij}}\left[\frac{P^2_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{y}_j\right] x_j y_j^4 \quad \text{(A.70)}$$

$$\frac{\partial \rho_{ij}}{\partial k^y_{05}} = -\frac{1}{\rho_{ij}}\left[\frac{P^2_i X^{mod}_j}{P^3_i X^{mod}_j} - \hat{y}_j\right] y_j^5 \quad \text{(A.71)}$$

The next component of the cost function is the set of constraints on the model correction equations. The cost elements that define these constraints are as follows:

$$Constr(\mathbf{x}) = m\sum_{j=1}^{n}\left(\kappa_{x_j}^2+\kappa_{y_j}^2+\kappa_{z_j}^2\right) \tag{A.72}$$

$$= \sum_{j=1}^{n}\chi_j^2 \qquad \text{where} \tag{A.73}$$

$$\chi_j = \sqrt{m\left(\kappa_{x_j}^2+\kappa_{y_j}^2+\kappa_{z_j}^2\right)} \tag{A.74}$$

As before, the optimizer accepts the χ_j in unsquared form. The jacobian entries for these cost elements are defined below:

$$\frac{\partial\chi_j}{\partial\kappa_{x_j}} = \frac{\kappa_{x_j}\sqrt{n}}{\sqrt{\kappa_{x_j}^2+\kappa_{y_j}^2+\kappa_{z_j}^2}} \tag{A.75}$$

$$\frac{\partial\chi_j}{\partial\kappa_{y_j}} = \frac{\kappa_{y_j}\sqrt{n}}{\sqrt{\kappa_{x_j}^2+\kappa_{y_j}^2+\kappa_{z_j}^2}} \tag{A.76}$$

$$\frac{\partial\chi_j}{\partial\kappa_{z_j}} = \frac{\kappa_{z_j}\sqrt{n}}{\sqrt{\kappa_{x_j}^2+\kappa_{y_j}^2+\kappa_{z_j}^2}} \tag{A.77}$$

The final component of the cost function is the set of weighted variances used to force shared parameters (denoted by the set Γ) to converge to a common value. These cost elements are described by the following equations:

$$Var(\mathbf{x}) = w_l\sum_{l\in\Gamma}\frac{\sum_{i=1}^{m}\left(p_{li}-\mu_l\right)^2}{m-1} \tag{A.78}$$

$$= \sum_{l\in\Gamma}\upsilon_l^2 \tag{A.79}$$

$$\upsilon_l = \sqrt{w_l\frac{\sum_{i=1}^{m}\left(p_{li}-\mu_l\right)^2}{m-1}} \tag{A.80}$$

The jacobian entries for the unsquared cost elements υ_l are shown below:

$$\frac{\partial\upsilon_l}{\partial p_{li}} = \sqrt{\frac{w_l}{\frac{\sum_{i=1}^{m}(p_{li}-\mu_l)^2}{m-1}}}\left(\frac{p_{li}-\mu_l}{m-1}\right) \tag{A.81}$$

Appendix B

KALMAN FILTERING FOR BALLOON TRACKING

This appendix provides an overview of Kalman filtering as it might be applied to intraoperative balloon tracking. While this is not a core part of this doctoral work, it was an active topic of discussion during this research and may therefore be helpful to others exploring related issues. This appendix does not introduce any signficant new material, but it will provide some insights into the application of Kalman filtering to intraoperative organ tracking.

Kalman filtering is commonly used to reduce the impact of random noise on an imaging or sensor system. Several filter models encompassing different motion characteristics have been developed and will be tested on the tracking data from the fluoroscope.

One of the challenges in any type of signal analysis is the management of ubiquitous error in the input data. For example, it is reasonable to assume that a parked car on the street will not suddenly move by three meters between sequential high-speed video frames, yet errors in a standard computer vision system could yield such a result. How can one formalize this type of knowledge to improve a tracking system? Kalman filtering is one popular approach to this problem.

B.1 Theory of Kalman Filters

The Kalman filter uses an iterative state model in which the current state, the current measurement values, and a statistical noise model are combined to predict the true state of the system at the next time step. In this case, 'predict the true state' means that the estimated optimal state minimizes the variance of the prediction error for systems that match the assumptions of the state model. Kalman introduced his state-based solution to the Wiener filter problem in [47–51], the first of which will serve as the basis for the discussion below.

B.1.1 Kalman's Process Model

Kalman defines his original ideal dynamic model to be:

$$x(t+1) = \Phi(t+1;t)\,x(t) + u(t) \tag{B.1}$$

$$y(t) = M(t)\,x(t) \tag{B.2}$$

with the following definitions:

$x(t)$	state vector of length n at time t
$\Phi(t+1;t)$	matrix describing transition to state $x(t+1)$ from state $x(t)$
$u(t)$	random process noise
$M(t)$	matrix mapping from states to measurements
$y(t)$	measurement vector at time t

This formulation was modified in [48] to include a transition matrix before $u(t)$ and a noise function $v(t)$ in $y(t)$. The new formulation was:

$$x(t+1) = \Phi(t+1;t)\,x(t) + G(t)\,u(t) \tag{B.3}$$

$$y(t) = M(t)\,x(t) + v(t) \tag{B.4}$$

The new formulation allows for more sophisticated compensation for higher-order errors in the process model (via $G(t)$) and for explicit modeling of the measurement noise via $v(t)$, as in [7].

B.1.2 State Prediction

Kalman's original paper stated the solution to the optimal estimation problem as a single-step process (although several calculations were involved). Because we want to use the filter's a priori predictions to direct our image search algorithms for target detection, Kalman's solutions is restated below in the form of a predictor-corrector approach. This formulation is described by Welch [106] as follows:

$$\hat{x}_k^- = A\,\hat{x}_{k-1} + B\,c_k \tag{B.5}$$

$$P_k^- = A\,P_{k-1}\,A^T + Q \tag{B.6}$$

The superscript-minus terms represent the a priori predictions of the state and prediction error covariance matrix respectively. The A term is the state transition matrix $\Phi(t+1;t)$ from

Kalman's formulation, c_k is a control input vector, B is a matrix mapping the control inputs into state changes, P_k is the covariance matrix of the residual error at time k, and Q is the covariance matrix for the random process noise function $u(t)$. The a priori prediction of state shown above can be used to drive the measurement function for the current time step.

B.1.3 Error Correction

After the measurements have been made for the current time step, the Kalman filter goes through a correction step that minimizes the residual error for future predictions. These error correction steps are as follows:

$$K_k = P_k^- H^T \left(H P_k^- H^T + R\right)^{-1} \tag{B.7}$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(z_k - H \hat{x}_k^-\right) \tag{B.8}$$

$$P_k = \left(I - K_k H\right) P_k^- \tag{B.9}$$

where K_k is the optimal Kalman gain at time k (similar to Kalman's original $D^*(t)$ function), H is the matrix that maps states to ideal measurements (identical to Kalman's $M(t)$ function), R is the covariance of the measurement noise function $v(t)$, and I is the identity matrix. The x_k and P_k results in this step represent the optimal estimates of the state and prediction error covariance for this time step. The x_k results from this calculation should be used for 3D reconstruction in the balloon-tracking task since they represent the best guess at the true state of the system.

B.1.4 Filter Performance and Error Analysis

The Kalman filter is a linear prediction filter, so the performance of the filter depends upon the linearity of the system under observation as compared to the state model as well as the noise characteristics of the system. For example, an object with significant accelerations could be tracked poorly by a state model that assumes constant position, while the same object could be tracked well by a system that assumes variable position and velocity.

Medical ventilators provide cyclic respiratory motion to patients under general anesthesia. The motion of the ventilators includes some periods of constant position, some with constant velocity and variable position, and some with constant acceleration but variable position and

velocity. If the pattern of the ventilator is indicative of the actual motion of the liver through the ventilation cycle, then a second-order model that handles variable positions and velocities may be necessary.

B.2 Motion Models for the TIPS application

The final step in the development of the theoretical basis for this tracking system is to adapt the Kalman filter methods to incorporate the position of centroid and orientation calculations as well as the reconstruction of image features into 3D. The following sections will address the selection of primary parameters to be estimated, the specification of a complete state model, and the incorporation of error modeling into the system.

B.2.1 Selection of Parameters

The first step in designing a filter is deciding on the data to be filtered. In this case, we have position and orientation measurements from the tracked balloon in the two fluoroscopic images. We can also reconstruct the balloon's three-dimensional position and orientation using these data. One must decide if the filtering should be performed on the raw image coordinate data prior to reconstruction (filter in image space), or if one should reconstruct the 3D position and orientation first and then filter those measurements (filter in world space).

The advantage of filtering the tracker data separately in each image is that the filter's noise reduction benefits would be applied directly to the source data. This could mean that within-image measurement errors would be reduced significantly before 3D reconstruction, thus reducing the error being propagated through the measurement chain.

The advantage in reconstructing the 3D position and two angles of orientation directly and filtering the 3D results is that the redundancy of the two image measurements would be used to reduce errors before the filter was applied. Six parameters would have to be filtered in the two-filter image-space model above, whereas only five parameters would be predicted in the world space model. This reduction in the filtering space is likely to yield more stable predictions and was chosen for the remainder of this work. The world-space model will be developed further in the following section.

B.2.2 Selection of State Model

As discussed in Section B.1.4, the performance of the Kalman filter depends upon the match between the process model and the actual process being measured. In particular, the order of the systems should be similar if accurate predictions are needed. Due to the uncertainty about the actual motion of the liver during the TIPS procedure, at least three models should be considered and tested for the task. In the first model, the true target will be assumed to have constant position and orientation and zero velocities for all parameters plus random measurement noise. This yields the following model:

$$X(t) = \begin{bmatrix} x \\ y \\ z \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} X(t-1) + u(t-1) \tag{B.10}$$

$$Y(t) = \begin{bmatrix} x \\ y \\ z \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} X(t) + v(t) \tag{B.11}$$

This model will work well if the liver is generally stationary.

The next model assumes zero acceleration, constant velocity, and variable positions plus

random noise for each parameter:

$$X(t) = \begin{bmatrix} x \\ x' \\ y \\ y' \\ z \\ z' \\ \alpha \\ \alpha' \\ \beta \\ \beta' \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} X(t-1) + u(t-1) \quad \text{(B.12)}$$

$$Y(t) = \begin{bmatrix} x \\ x' \\ y \\ y' \\ z \\ z' \\ \alpha \\ \alpha' \\ \beta \\ \beta' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} X(t) + v(t) \quad \text{(B.13)}$$

The primed entries indicated derivatives (velocities), while the non-primed entries represent positions.

The final model assumes constant acceleration with variable velocity and position. For the sake of brevity, the model will not be presented here. It is analogous to the velocity model except that the current acceleration is also stored and velocity and acceleration are factored into position and velocity estimates.

Each of these models should be trained and tested with data from real patients to determine which is most effective.

B.2.3 Adaptive Region Searching Driven by Kalman Filter Residuals

The Kalman filter residuals provide a measure of the uncertainty of current estimations, and thus could be used to expand or contract the object search region as needed. This concept is discussed in [108].

Appendix C

OTHER SUPPORTING WORK

This appendix reviews an assortment of material related to the design of an intraoperative guidance system. Topics include vessel modeling and one example of an operative guidance system that has been shown to have utility in a clinical task. These sections do not introduce any new research; they are solely provided as repositories of references and ideas for others who may pursue similar research.

C.1 Vessel Modeling

The extraction of blood vessels from pre-operative imagery has been described extensively [1–3] and the accuracy of its output has been evaluated [12]. Aylward's techniques extract vessels by traversing the intensity height ridges in a 3D image set. Vessel scale is determined using a sub-voxel tubule detection filter. The 3D/2D registration of a vascular model to intraoperative angiograms was developed and evaluated for accuracy in Liu's dissertation and published works [13, 57, 58]. Reconstruction of 3D curves (e.g., needles) has also been demonstrated and evaluated [14, 15]. These components will be further evaluated as part of the TIPS system development, but that work is not a part of this dissertation.

C.2 Clinical Utility of 3D Visualization

This project is predicated on the belief that 3D visualization could have a meaningful place in clinical practice. While this is intuitively sensible, it is also valuable to have experimental data supporting such a belief. The following section describes an experiment conducted by our research group to test the impact of 3D visualization on biopsy placement accuracy. This text is adapted from our publications on the topic [81, 83]. The goal of this section is not to prove that our approach to TIPS guidance will be effective, but rather to show that 3D guidance can be of value in at least one relevant context.

C.2.1 Introduction

Our research group at the University of North Carolina has been working in the area of augmented reality (AR) visualization for ultrasound examinations and ultrasound-guided procedures for nearly a decade [5,32,34,45,92–94]. The vision for this project is to allow physicians to directly see into a patient, aided by real-time computer graphics and augmented reality technology. The notion of augmenting the view of one's surroundings with computer-generated images has its roots in Ivan Sutherland's seminal paper [97], which described a system with a head-mounted display (HMD) whose synthetic images the user could see optically overlaid on the view of the room around him. Many years of research, both in the general AR field [4] as well as in specific medical AR applications [24, 62, 96], have resulted in considerable improvement in each of the key technologies.

Using our biopsy guidance system in January 1996, a trained physician (Pisano) was able to guide a needle into a lesion within an artificial breast training phantom and report that the task was "easy" (Figure C.1). A subsequent test with a human subject progressed to where the needle was partially inserted towards the target lesion, at which point the physician was forced to abandon the AR guidance and continue the intervention with conventional ultrasound guidance technology. During this and several subsequent experiments it slowly became clear that despite the technological advancements effective patient studies were still not possible. This was mostly due to cumbersome equipment and inadequate tracking technology [32].

We have spent the intervening years developing an enhanced guidance system, which is now being used in live patient studies. In the following sections, we describe the new developments in our guidance system. We also describe the design and report the results of a randomized, controlled study to determine the relative effectiveness of our new AR system versus traditional ultrasound. We conclude with a description of our current and future work.

C.2.2 Materials and Methods

Other papers have described our system design in detail [32, 34, 45, 92, 94]. In the following sections, we provide a brief overview of our current system and the full design of our recent biopsy accuracy experiment.

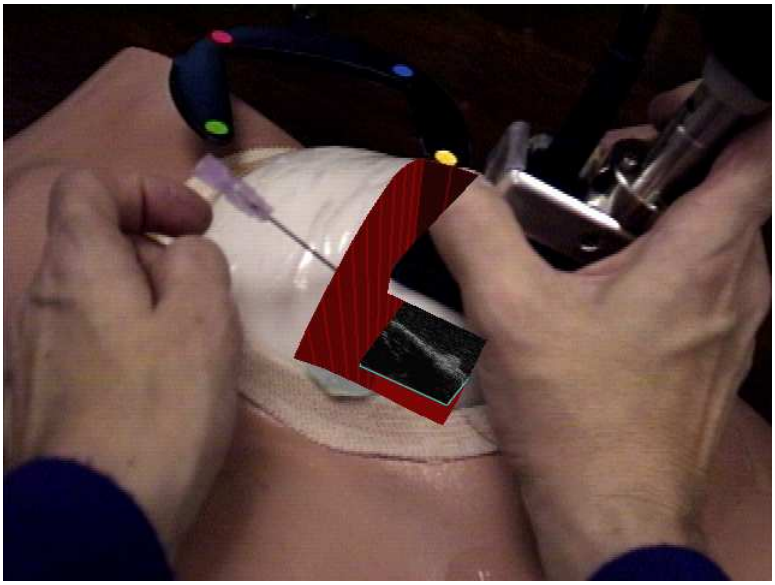


Figure C.1: View from the head-mounted display during a 1996 biopsy guidance experiment. The physician has inserted a cyst aspiration needle into a lesion within a breast phantom and holds the ultrasound transducer in her right hand. Correct ultrasound probe calibration and accurate tracking yield alignment between real needle and image of the needle in the ultrasound slice. The colored dots in the background are fiducials for head tracking correction (not used in our current system).

Augmented Reality Guidance System

System Overview. Our AR guidance system consists of four major components: an ultrasound imaging system, an instrument tracking system, a graphics and computation platform, and a head-mounted display (HMD). The ultrasound imaging system produces a live video stream that is digitized and fed into our computation platform. The computation platform also receives data on the position and orientation of the ultrasound scanner, the HMD, and the biopsy needle from our tracking system. This information is combined with our computational models of the instruments and patient to produce a "virtual" procedure room. Selected features of this room are overlaid on the live video from our video-see-through HMD and fed back into the HMD, providing the user with an augmented view of the real world. Additional details of our system, including our modeling and calibration methods, can be found in the references cited above.

Ultrasound Imaging System. We are currently using a PIE Medical Ultrasound Scanner 350 to acquire ultrasound images during our experiments. The measurements on our ultrasound scanner have been tested to be within the manufacturer's specified error of less than 1mm over a 20mm span. The geometric transform between the ultrasound scanner images and our computational model is determined using a custom calibration phantom of interlaced silk threads.

Tracking System. We use an Image-Guided Technologies FlashPoint 5000 opto-electronic tracker in our system. The HMD, the ultrasound probe and the biopsy needle are all equipped with infrared LEDs. The FlashPoint delivers readings of the positions of these LEDs to the graphics computer with sub-millimeter accuracy. This HMD tracking technology is not quite as accurate as the closed-loop method used in our original 1996 system [94], but it is superior to magnetic technologies and does not encumber the user's field of view (and the sterile operating field) with fiducials. The ultrasound probe is also tracked opto-electronically. It uses a specially developed 9-LED device that allows rotations up to 80 to any side without losing acquisition, thus freeing the physician to position and orient the probe in the most adequate way for a particular intervention.

Figure C.2: Head-mounted display view during this phantom biopsy experiment (in 2000). Both the ultrasound probe (left hand) and the biopsy needle (right hand) are tracked. The needle aims at the bright lesion visible in the ultrasound slice. The system displays the projection of the needle onto the plane of the ultrasound slice (blue lines) and also displays the projected trajectory of the needle if it were fired at this moment (yellow markers)

Graphics and Computation Platform. The system runs on an SGI Onyx2 Reality Monster graphics computer equipped with four DIVO digital video input/output boards, allowing simultaneous capture of eight video streams. The software runs at frame rates of 20-30 Hz in stereo on this platform. Figure C.2 shows imagery displayed by our system during an experiment with a breast training phantom in late 2000.

Head-Mounted Display. We have modified a stereoscopic Sony Glasstron LDI-D100 HMD for use as our display system. This HMD provides full color, stereo, SVGA (800x600) resolution displays in a lightweight design. We have added an aluminum superstructure to hold two Toshiba IK-SM43H video cameras for image capture and three infrared LEDs for opto-electronic tracking. Figure C.3 shows the model of our HMD used in this experiment; a more recent version of the HMD has been described by State [92, 95]. This "video see-through" [4] device and its operation are described in detail in [92].

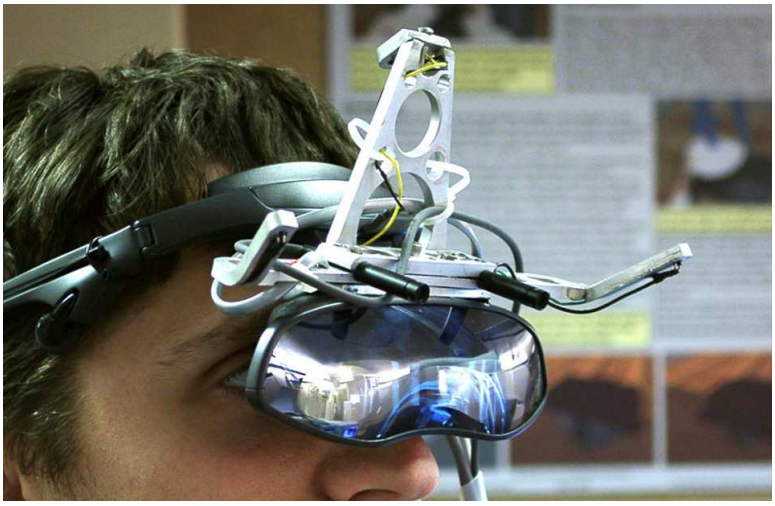


Figure C.3: Video-see-through augmented reality head-mounted display (HMD) built on the basis of a Sony Glasstron LDI-D100 device. The aluminum superstructure holds two miniature video cameras for image capture and three infrared LEDs for opto-electronic tracking of the HMD.

Design of Biopsy Guidance Study

We have performed an experiment to compare our AR guidance system to standard ultrasound guidance for the task of targeting needle biopsies in training phantoms. Our hypothesis was that the two guidance methods would be comparable in terms of needle placement accuracy for this task, indicating that it is safe to evaluate the AR system in humans.

The experimental component of this study was performed using the AR system described above. The control component was performed using only the PIE Medical Ultrasound Scanner 350 component of our system without any computer augmentation.

Biopsy Task. Our task under evaluation was diagnostic biopsy of a simulated solid breast mass. Standard ultrasound training phantoms (Model 52 Biopsy Phantom, Computerized Imaging Reference Systems, Inc., Norfolk, VA) were used as our biopsy subjects. These phantoms each contained six tumor-like targets placed randomly throughout an ultrasound-compatible gel mold. The phantoms are approximately the size and shape of an average human breast. A new phantom was used whenever the radiologist felt that artifacts from previous

biopsies were interfering with the current task.

Biopsies were performed using the standard procedure for breast biopsies in the US. In this approach, the lesion is located using the ultrasound scanner, and then the core biopsy needle is inserted in line with the plane of the ultrasound image. As the needle is advanced, the ultrasound probe is moved to keep the needle in full view at all times. To confirm the out-of-plane position, the needle is intermittently held steady while the ultrasound scanner is turned 90 degrees and the needle's lateral placement is determined. This process is repeated until the biopsy needle is within range of the lesion. With the needle held in the plane of the ultrasound scanner, the biopsy mechanism is fired and the resulting biopsy position is confirmed in the standard plane and in the perpendicular ultrasound plane. This process is repeated to take biopsies from the center of the lesion and at the three, six, nine, and twelve o'clock positions around the perimeter of the lesion (as viewed on the plane orthogonal to the axis of the biopsy needle). Biopsy techniques are discussed in additional detail in [72, 75].

For each selected lesion in a phantom, five biopsies were performed using the method described above with a 14-gauge Monopty core biopsy needle (C. R. Bard, Inc., Covington, GA). The needle was withdrawn from the phantom after each biopsy attempt. The ultrasound video from each biopsy was reformatted and recorded directly from the ultrasound scanner to DV tape for later evaluation. The biopsies were all performed during a single experimental session.

Randomization and Control Scheme. This study was designed as a randomized, controlled trial to limit the effects of confounding factors. A single board-certified radiologist (Pisano) performed all of the biopsies in this experiment (Figure C.4). Ten targets within the phantoms were sequentially selected; five biopsies were performed on each lesion before selecting the next target. Randomization to the two guidance methods was performed by a coin flip before the selection of each biopsy target.

Evaluation of Accuracy. Another board-certified radiologist (Cherie Kuzmiak, DO) evaluated the ultrasound video to determine the accuracy of each biopsy. The evaluator was blinded to the method of guidance for each biopsy. For each biopsy, she determined the geometric distance (in mm) between the ideal biopsy target point and the actual biopsy positions in the plane orthogonal to the needle. The evaluator also measured the dimensions

Figure C.4: Lab view (top) and ultrasound image (bottom) while the physician, wearing the Glasstron-based head-mounted display (HMD), performs a controlled study with the 2000 system. She holds the opto-electronically tracked ultrasound probe and biopsy needle in her left and right hands, respectively. Refer to Figure 3 for an example of the view from the HMD. The horizontal streaks in the ultrasound image are common artifacts of acoustic reflection.

of the lesions along the needle axis and along two perpendicular directions (approximately vertical and horizontal). The results were later entered into an Excel spreadsheet by another individual (MR) and associated with the corresponding guidance method.

The geometric distances mentioned above were measured on an NTSC display with respect to the ultrasound machine's displayed reference ruler. This is the standard clinical technique for anatomical measurements in breast biopsy. While it would be desirable to have confirmation from an independent modality, we know of no alternative "gold standard" for accurately validating these measurements. We believe that the randomization and blinding scheme should minimize the impact of this uncertainty by minimizing the likelihood of a bias in measurement error in favor of either method.

Statistical Analyses. Descriptive statistics (mean, std) of the error distances were calculated. Separate and combined results were computed for the HMD and standard ultrasound CRT display methods for each location, mean error across locations and the mean of the maximum lesion dimension. The primary analysis was a repeated measures analysis of variance (REPM ANOVA) utilized to address the multiple locations targeted within each lesion (a within-"subject" repeated measures dimension). The SAS procedure GLM was utilized.

To rule out lesion size bias as contributing to the effect attributed to display method in the primary analysis, we performed an exploratory full model in every cell (FMIC) REPM ANOVA analysis to show that the effect due to lesion size was not significant between the display methods. The FMIC was then reduced to a multivariate analysis of covariance (MANCOVA) model and reanalyzed. Maximum lesion dimension (in mm) was the measure we chose to represent lesion size.

C.2.3 Results

A total of fifty biopsies were performed: twenty-five in each of the AR guidance and standard guidance groups. The mean error distances for each of these groups are shown in Table 1 below, while Figure C.5 shows the distribution of errors for both groups. A repeated measures analysis of variance indicated that the HMD display method led to a statistically significantly smaller mean deviation from the desired target than did the CRT display method. (2.48mm for control versus 1.62mm for augmented reality, $p < 0.02$). The biopsy location and the

(All measures in mm, mean ± std dev)	Standard Guidance	AR Guidance
Error at Center	4.20 ± 1.92	1.50 ± 1.41
Error at 3 O'clock	2.00 ± 1.87	1.70 ± 0.67
Error at 6 O'clock	1.20 ± 0.84	0.90 ± 1.02
Error at 9 O'clock	2.00 ± 1.58	0.80 ± 1.30
Error at 12 O'clock	3.00 ± 2.00	3.20 ± 2.05
Mean Error across Locations	2.48 ± 0.44	1.62 ± 0.48
Mean of Maximum Lesion Dimension	10.50 ± 3.26	12.00 ± 2.09

Table C.1: Results from the phantom biopsy study

location-display combination did not yield statistically significant effects upon the accuracy.

The supportive FMIC ANOVA and MANCOVA analyses of the effects of lesion dimensions upon accuracy indicated that the maximum lesion dimension had no significant effect upon placement error ($p > 0.05$ for the main effect and all combinations involving maximum lesion dimension). These results indicate that the guidance method was the only factor associated with a statistically significant difference in placement error.

The measured lesion dimensions were consistent with the manufacturer's specifications for the biopsy phantoms. The lesions were stated to be 6-12mm in diameter by the manufacturer; our measurements ranged from 5mm to 12.5mm with a mean of 10.3mm.

C.2.4 Conclusions

The results of the above study indicate that the AR guidance system yielded statistically significantly improved accuracy as compared to the standard ultrasound guidance method for our expert user. In fact, we did not expect the AR system to be as good as the conventional guidance technique, especially for the expert user (Pisano). Our goal was to demonstrate the system's effectiveness on a procedure that is simple and not dangerous to the patient. The indication that the AR technique may be better in this comparison, where the advantage should go to the conventional approach, is both surprising and encouraging. We are hopeful that user studies with less experienced physicians may show an even greater improvement

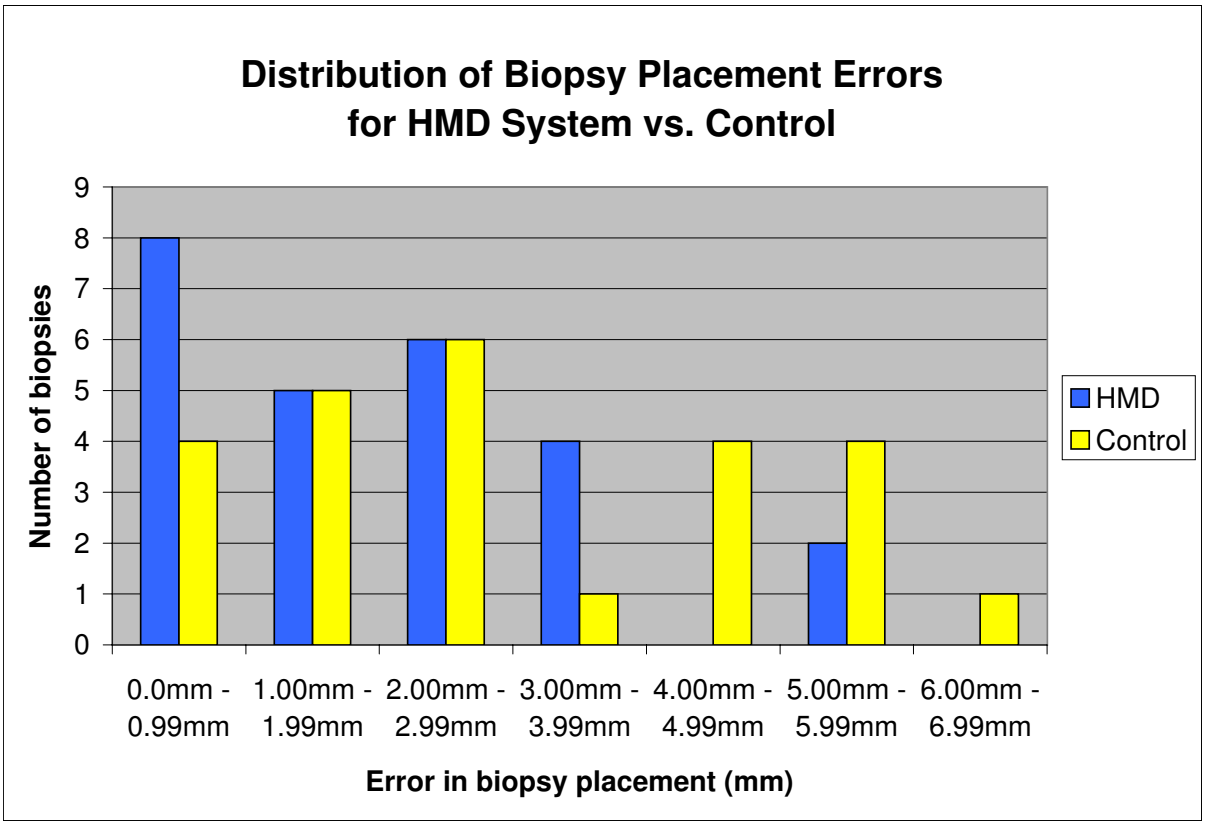


Figure C.5: Distribution of biopsy placement errors using head-mounted display (HMD; dark bars) and standard ultrasound guidance methods (Control; light bars). The HMD group showed a statistically significant reduction in mean biopsy placement error (1.62mm for HMD versus 2.48mm for control, $p < 0.02$).

using an AR approach.

INDEX