An Analysis of Tracking Error in Image Guided Neurosurgery

Ian J. Gerard, B.Sc.

Medical Physics Unit
McGill University, Montreal

August 2013

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Each year, thousands of Canadians undergo neurosurgery to areas of the brain that are critical to movement, vision, sensation, or language. Recent literature demonstrates a significantly increased survival benefit with complete resection of primary and secondary brain tumours. Image guided neurosurgery uses technology that tracks a patient and special tools simultaneously to use preoperative images in order to help guide a surgeon through the surgery. There are many different types of errors that can arise during these types of interventions related to technical, physical and biological factors. The aim of this study was to quantify some of the technical and physical factors that contribute to error in these interventions. Errors associated with tracking, tool calibration and registration between a physical object and a corresponding image were all investigated and compared to theoretical descriptions of these errors. A precision milled linear testing apparatus was constructed to perform the bulk of the measurements which were broken into three categories: the fiducial localization errors (FLE), the fiducial registration errors (FRE) and the target registration errors (TRE). The fiducial localization errors deal with errors in locating a physical point and correspond to the tracking error and tool calibration errors. The tracking error (jitter) was measured as the camera's ability to consistently report the proper location of a tool and was shown to increase in a quadratic fashion with distance normal to the camera and the jitter ranged between 0.15 mm – 0.6 mm. The tool calibration error was measured and showed to increase as a function of distance from the camera as well as distance from the reference tool with calibrations ranging from 0.2 mm – 0.7 mm for the NDI pointer and 0.2 mm – 0.8 mm for the Traxtal pointer. The fiducial registration error was investigated by registering a custom built Linear Testing Apparatus (LTA) to a corresponding image volume. The FRE was shown to improve when more points were used up until a plateau value was reached which corresponded to the total FLE. This value was on the order of 0.8 mm. The distributions of the TRE were investigated for four camera-pointer pairs and two fiducial configurations and were shown to follow a chi squared distribution with the highest error generally around fiducial points and the highest variation of the TRE also around the fiducial points. Most of the results observed in this work agreed well with the previously established theory.

Résumé

Chaque année, des milliers de Canadiens subissent une intervention chirurgicale à proximité de zones du cerveau qui sont essentielles au mouvement, à la vision, à la sensation ou au langage. Cela crée deux contraintes contradictoires qui doivent être équilibrées pendant la chirurgie pour chaque patient. Il faut maximiser la résection de la lésion tout en minimisant le déficit neurologique pour le patient. La neurochirurgie guidée par l'image utilise une technologie qui permet de suivre simultanément la position d'outils spéciaux et du patient pour guider le chirurgien sur des images préopératoires durant la chirurgie. Il y a plusieurs types d'erreurs qui peuvent survenir lors de ce type d'intervention. Ces erreurs sont liées à des facteurs techniques, physiques et biologiques. Le but de cette étude est de quantifier certains des facteurs techniques et physiques qui contribuent à l'imprécision de ces interventions. Les erreurs liées au suivi des outils, au calibrage des outils et au recallage entre un objet physique et une image correspondante ont tous été étudiées et comparées à leurs descriptions théoriques. Un montage de test linéaire (LTA) a été usiné avec précision pour effectuer la majeure partie des mesures qui ont été divisés en trois catégories: les FLE, les FRE et les TRE. Les FLE sont des erreurs de la localisation d'un point de vue physique et correspondent à des erreurs de suivi et des erreurs de calibrage des outils. L'erreur de suivi d'un outil (jitter) a été mesurée comme étant la capacité de l'appareil à rapporter systématiquement la position exacte d'un outil et il a été démontré qu'il augmente de façon quadratique avec la distance le long de l'axe de la caméra et que sa valeur se situe entre 0,15 mm et 0,6 mm. L'erreur de calibrage des outils a été mesurée et il a été démontré qu'elle augmente en fonction de la distance de la caméra ainsi que de la distance à de l'outil de référence avec une erreur de calibration allant de 0,2 mm - 0,7 mm pour le pointeur NDI et 0,2 mm - 0,8 mm pour le pointeur Traxtal. Le FRE a été étudié en recallant la LTA sur une image volumétrique correspondante. Il a été démontré que le FRE s'améliorer avec le nombre de points utilisés jusqu'à ce qu'une valeur de seuil ait été atteinte, valeur qui correspond au FLE total. Cette valeur est de l'ordre de 0,8 mm. Les distributions du TRE ont été étudiés pour quatre paires de caméra-pointeur et deux configurations de points et il a été démontré qu'elles suivent une distribution chi carré avec la plus grande erreur généralement autour des points de référence et la variation la plus élevée de la TRE également autour des points de référence. La plupart des résultats obtenus pour cet étude concordent avec la théorie établie précédemment.

Acknowledgements

I would like to thank:

Dr. D. Louis Collins, my supervisor, for his guidance and assistance and his continuous constructive criticism on all work presented to him. Being in your lab for the last year has been very enjoyable and I look forward to continuing on with my doctoral work under your supervision.

All the members of the Image Processing Laboratory and Brain Imaging Center, in particular: Anka Kochanowska, Simon Drouin, and Yiming Xiao who's help with programming and customizing IBIS were very useful for much of the work that was and will be done in the future.

Joe Larkin, the machinist at the Montreal General Hospital, who's expertise in precision manufacturing helped create my linear testing apparatus where most of the measurements in this thesis were performed.

Dr. Jeffery Hall, a collaborative neurosurgeon, who has helped me understand some of the aspects of where neuronavigation is needed to improve in surgery and who introduced me to the clinical side of image guided neurosurgery and will be a large partner in my doctoral work.

Kelvin Mok, head of clinical neuronavigation at the MNI, who was always available to offer advice and to lend special tools and equipment that could be used to perform measurements.

Dr. Bruce Pike, my professor for the Medical Imaging course, who introduced me to Louis and the MNI thus opening the door for me to be part of the great MNI community.

My mother, who has always been supportive through all of my academic endeavours.

Thank you all.

Table of Contents

AUTHOR'S DECLARATION ii
Abstract iii
Résumé iv
Acknowledgements v
Table of Contents vi
List of Figures viii
List of Tables ix
Chapter 1 Introduction 10
Chapter 2 Background 13
2.1 Preoperative Imaging 13
2.1.1 Image Guided Neurosurgery Based on Preoperative MR images 14
2.2 Neuronavigation Systems 15
2.2.1 Relating Coordinate Systems 17
2.2.2 Tracked Pointer Calibration 18
2.2.3 Main Sources of Error 19
2.2.3.1 Pointer Calibration Errors 19
2.2.3.2 Tracking Errors 20
2.2.3.3 Patient-to-Image Registration Errors 20
2.2.3.4 Preoperative Imaging Acquisition Errors 20
2.2.3.5 Brain Shift 21
2.3 Image Registration 21
2.3.1 Registration Framework 22
2.3.2 Steps of Image Registration 23
2.4 Registration Errors 23
2.4.1 Fiducial Localization Error (FLE) 24
2.4.2 Fiducial Registration Error (FRE) 25
2.4.3 Target Registration Error (TRE) 27
2.4.4 An Illustrative Analogy 28
2.4.5 Relationships Between FLE, FRE, and TRE 29
2.5 Previous work in IGNS error analysis 31
Chapter 3 Methods 33

3.1 Materials 33
3.1.1 Linear Testing Apparatus 35
3.2 Pointer Calibration Error 38
3.3 Tracking Error Quantification 39
3.4 Error Distribution in the LTA Volume 41
Chapter 4 Results 43
4.1 Pointer Calibration Errors 43
4.2 Tracking Error Measurements 43
4.3 Fiducial Registration Error Distributions in the LTA 47
4.4 Target Registration Error Distribution in the LTA 51
Chapter 5 Discussion 56
5.1 Pointer Calibration Errors 56
5.2 Tracking Error 57
5.3 Registration Error Distributions in the LTA 58
5.4 Other Sources of Error 60
Chapter 6 Conclusion and Future Work 62
Bibliography 62

List of Figures

Figure 2.1 Preoperative MR image set of a patient undergoing neurosurgery 14
Figure 2.2 Typical structure surfaces that would be seen in a neuronavigation environment 15
Figure 2.3 Typical setup of neuronavigation unit in the operating room of the MNH. 16
Figure 2.4 Schematic diagram showing the relationship between different coordinate systems 17
Figure 2.5 Basic components and workflow of a registration framework 22
Figure 2.6 Various types of registration error 24
Figure 2.7 The horizontal axis is TRE in units of mean TRE with the vertical line at the mean. 28
Figure 3.1 All the trackers and tracked tools used for measurements in the described experiments. . 34
Figure 3.2 A schematic diagram showing the cylindrical measurement volume of the Polaris. 35
Figure 3.3 Left: view of the LTA. Middle: precision milled spacing rods 36
Figure 3.4 Left: 2 identical pieces of the LTA with the alignment piece before assembly.. 37
Figure 3.5 Schematic diagram of the setup for pointer calibration experiment one and two. 39
Figure 3.6 Left: The relative position of the camera and the LTA plate 40
Figure 3.7 Overhead view of camera setup for registration error experiments. 42
Figure 4.1 Pointer calibration results for both Polaris Cameras. 45
Figure 4.2 Jitter as a function of depth for four different camera-marker combinations 46
Figure 4.3 Max jitter values for the lowest 90% of the jitter data 46
Figure 4.4 Boxplot showing the ranges of the individual FREs. 48
Figure 4.5 Box plot showing the standard deviation of the individual FREs. 48
Figure 4.6 Evolution of FRE as a function of number of points used for Camera 1 49
Figure 4.7 Evolution of FRE as a function of number of points used for Camera 2 50
Figure 4.8 Histogram of the TRE value distributions in terms of the mean TRE. 52
Figure 4.9 Illustrative example of the mean and standard deviation contours for a slice 52
Figure 4.10 Histogram of the TRE value distributions in terms of the mean TRE for the. 53
Figure 4.11 Illustrative example of the mean and standard deviation contours for a slice 53
Figure 4.12 Histogram of the TRE value distributions in terms of the mean TRE for 54
Figure 4.13 Illustrative example of the mean and standard deviation contours for a slice 54
Figure 4.14 Histogram of the TRE value distributions in terms of the mean TRE for 55
Figure 4.15 Illustrative example of the mean and standard deviation contours for a slice 55

List of Tables

Table 2.1 Patient-Image Registration Accuracy in the Literature 32
Table 3.1 Camera Positions 42
Table 4.1 Camera-Marker Jitter Value Ranges..... 45

Chapter 1

Introduction

Each year, thousands of Canadians undergo neurosurgery for resection of lesions such as tumours, epileptic foci and vascular malformations in close proximity to areas of the brain that are critical to movement, vision, sensation, or language. Recent literature demonstrates a significantly increased survival benefit with complete resection of primary and secondary brain tumours (McGirt, et al. 2009). This creates two competing constraints that must be balanced during surgery for each patient: *to achieve maximal resection of the lesion while causing minimal neurological deficit.* For many neurosurgeries, preoperative planning is done on an image guided neurosurgery (IGNS) system that gives a surgeon the tools to visualize, interpret and navigate through patient specific volumes of anatomical, vascular and functional information and also gives them the opportunity to investigate some of their inter-relationships. For example, a tumour and the surrounding anatomy can be viewed in exquisite detail with magnetic resonance imaging (MRI), eloquent cortex and critical pathways adjacent to the tumour that must remain intact can be identified by positron emission tomography (PET) and functional MRI (fMRI), a technique called digital subtraction angiography (DSA) is used to visualize and plan a vessel-free path to a surgical target and diffusion tensor imaging (DTI) is used to minimize damage to white-matter fibre tracks. Combining all or a subset of these images help surgeons evaluate the risks associated with a specific procedure, to determine the most appropriate surgical strategy to remove the lesion of interest and to plan the procedure with incredible detail. During surgery, the IGNS system is used for surgical guidance within a neuronavigation environment that tracks the patient as well as a set of specialized surgical tools with an image-to-patient mapping. All of these systems relate the real-world coordinates of a patient to those of the preoperative images

using a rigid body transformation. This transformation is determined by identifying implanted markers or external features on both the patient and the images. This allows a surgeon to point to a specific location or physical landmark on the patient in the operating room and see the corresponding anatomy in the preoperative images. There are many advantages associated with these systems, the main one being greater positional accuracy during the surgical intervention that results in numerous benefits. Perhaps more importantly, these systems enable surgery of very small lesions and previously inoperable cases due to the identification of surgical corridors and strategies through IGNS identified non-critical areas.

Unfortunately, IGNS systems have not lived up to their promise or potential. There are many sources of error associated with IGNS systems, all of which lead to the position of a tracked probe being incorrectly reported in relation to the images being used for guidance. Many of these errors arise from the failure of two basic assumptions:

(1) That the equipment, registration and images are perfectly accurate. This assumes that the pointer tracking device is free of any positioning error, the mapping between patient and image is perfect and that the preoperative images are free from spatial distortion.

(2) That the equipment and volume of interest form a rigid system. This assumes the structures of interest remain in the same position with respect to the originally picked external points throughout the entire procedure.

Most commercially available tracking devices can provide position and orientation information with sub-millimeter accuracy (Khadem, et al. 2000) and point matching between homologous point pairs manually identified on the patient and images can yield a mapping with accuracy on the order of 2 – 3 mm (Cuchet, et al. 1995). Image distortion in MRI (the most common image modality used in IGNS) is highly dependent on acquisition parameters and geometric distortion can be on the order of 2 – 3 mm if precautions are not taken to avoid it (C. R. Maurer Jr., et al. 1996). During surgery, the movement of brain tissue invalidates the patient-to-image mapping, reducing the effectiveness and reliability of the IGNS system since these changes are not portrayed in the preoperative images that are being used for surgical guidance. This movement, referred here forth as “brain shift”, is a complex spatiotemporal phenomenon that is influenced by physiological, chemical and physical factors. Commercial IGNS rely solely on preoperative data and thus cannot properly account for this type of deformation which has been shown to range from 2 – 30 mm (Hill, et al. 1998), (Roberts, et al. 1998). As a result of these sources of error, surgeons at most hospitals use the IGNS system as a tool to

approach a surgical target, but justifiably no longer trust the system during the resection for many cases.

In this thesis, errors associated with the IGNS tracking technology and tracking tools are investigated. The effect of the tracking as well as its impact on different forms of registration errors is evaluated for many different experimental setups; specifically those analogous to the way equipment would be placed and procedures would be performed in a clinical scenario. The goal is to minimize the error associated with tracking and its impact on the registration in order to improve the accuracy and reliability of IGNS systems.

Chapter 2

Background

This chapter provides an introduction to various topics covered in the work of this thesis. It begins with an overview of IGNS and the different tools incorporated in these types of interventions. It is followed by highlighting different sources of error that are introduced by each of the different components as well as a short review of some of the previous work done in IGNS accuracy analysis. When presenting equations to describe specific phenomena, the following notation is used: capital, bold, italicized characters are used to represent vectors and capital italicized characters represent matrices. For example: $\boldsymbol{P}$ is a vector and P is a matrix.

2.1 Preoperative Imaging

In IGNS, MRI is the most common imaging modality obtained as preoperative data for planning and guidance of neurosurgeries and will be the focus of this section. However, it is worth noting that computed tomography (CT) images are still used for certain neurosurgery cases. The first application of MRI was published in 1973, by Paul Lauterbur (Lauterbur 1973). Two test tubes of heavy water were placed in a sample of normal water. A two-dimensional image was created by applying magnetic field gradients to a conventional nuclear magnetic resonance (NMR) system. Since its birth in 1973, MRI has become significantly advanced and improved upon and is now used in various medical fields for visualizing anatomical structures as well as for functional and metabolic information. An important characteristic of MR imaging is the vast flexibility the modality offers. The wide range of physical parameters to image as well as the plethora of instrumental parameters to set for control of image contrast opens the possibilities of noninvasive *in vivo* physiological studies. Patients who are candidates for neurosurgery will generally undergo an initial MRI exam before

surgery involving a series of different types of scans intended to highlight or suppress certain features of interest. An example of a set of clinical MR images for a patient that is undergoing neurosurgery for a tumour resection is shown in Figure 2.1.

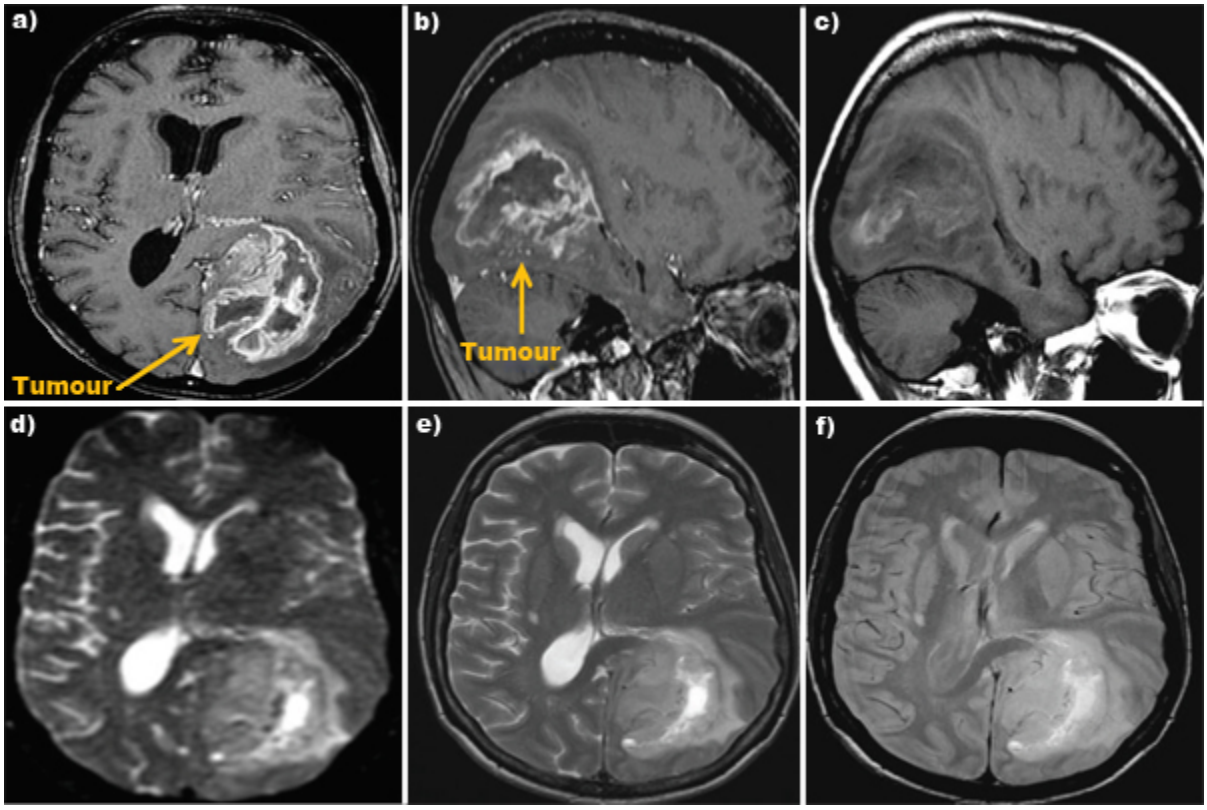

Figure 2.1 Preoperative MR image set of a patient undergoing neurosurgery including an axial (a) and sagittal (b) view of T1 weighted image with gadolinium enhancement, a T1 fluid attenuation inversion recovery (FLAIR) image (c), a diffusion weighted echo-planar image (DW EPI) (d), a T2 weighted image (e) and a proton density (PD) image (f). [Images reproduced from (Mercier 2011)]

2.1.1 Image Guided Neurosurgery Based on Preoperative MR images

In addition to being used as a diagnostic tool, MR images are used to plan the surgical intervention and serve as the basis for guidance within a neuronavigation environment. During preoperative planning, IGNS systems provide the surgeon with tools to visualize and interpret patient specific volumes of anatomical, vascular and functional data as well as determining some of their interrelationships. Together, these tools allow a surgeon to evaluate the different risks involved with a specific intervention and allow for the most optimal surgical strategy for the procedure.

Neuronavigation is a technique that uses preoperative images, a tracking device and special tracked surgical tools to help guide a surgery. A patient-to-image mapping is first determined to relate the physical position of the patient with the position of the images. The surgeon can thus point to a specific location on a patient and see the corresponding anatomy in the preoperative images. Neuronavigation and the patient-to-image mapping are discussed in sections 2.2 and 2.2.1. When eloquent areas of the brain are involved in the surgical region, functional imaging can be performed to determine the extremities of these regions. These areas can then be segmented, turned into 3-D surfaces and overlaid on top of the preoperative images for easier visualization. Some of the structures that are generally segmented include: the tumour(s), main blood vessels, cortex and the skin surface. An example of how these segmented volumes and surfaces generally appear in the neuronavigation system is shown in Figure 2.2.

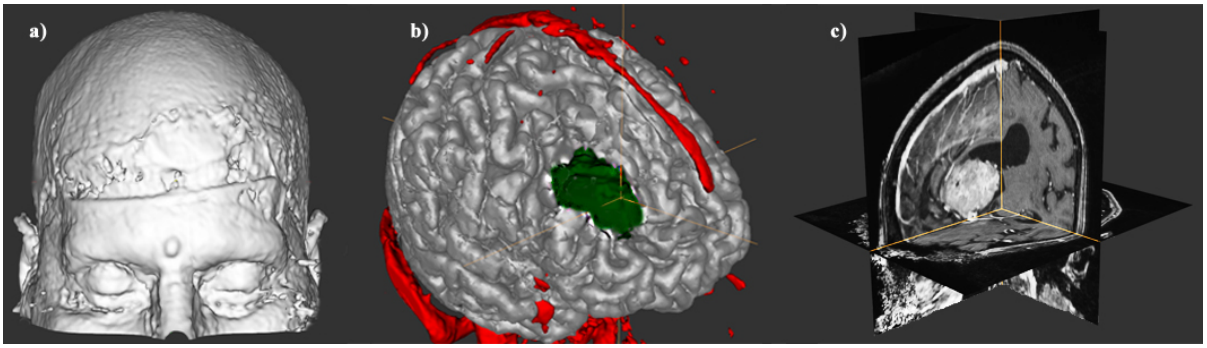

Figure 2.2 Typical structure surfaces that would be seen in a neuronavigation environment for an IGNS intervention including (a) the skin, (b) the cortex [grey], main blood vessels [red], and the tumour [green]. In (b) a portion of the cortex has been hidden for easy viewing of the tumour surface. In (c) there is an example of the corresponding preoperative MRI slices arranged in a 3D configuration.

The combination of imaging, planning and image guidance during surgery improves the positional accuracy of the intervention that gives rise to many benefits including: (i) minimally invasive cranial openings, (ii) accurate localization of sub-cortical lesions, (iii) reduction in blood loss, (iv) reduction in operating time, (v) avoidance of eloquent tissue and critical pathways, and (vi) a decrease in complication rate and thus reduction in intensive care unit and hospital stays.

2.2 Neuronavigation Systems

Neuronavigation is a term that is used to describe the ensemble of tools used for IGNS. All neuronavigation systems consist of several main components organized in a particular way. First, a

tracking system is used. This is commonly in the form of an infrared camera, but in some cases magnetic (Sagi, et al. 2003) or ultrasonic tracking (O'Donnell., et al. 1994) can be used. Since Polaris stereo infrared cameras (Northern Digital Instruments, Waterloo, Canada) are used at the Montreal Neurological Institute and Hospital (MNI/MNH) for tracking, focus will be placed on this technology only. In addition to a tracking camera, a reference tool and surgical pointing tool with special reflective spheres are used. The reference defines the origin of the camera, or world space, and the tracked pointer is used to navigate within this space. The Polaris camera uses stereo triangulation to locate the reflective spheres placed on the both the reference and pointer. Finally, a navigation console is needed to display the preoperative images and 3D surfaces. The console is also connected to the camera in order to display the position of the pointer tool relative to the preoperative images. A typical operating room setup for an IGNS case is shown in Figure 2.3.

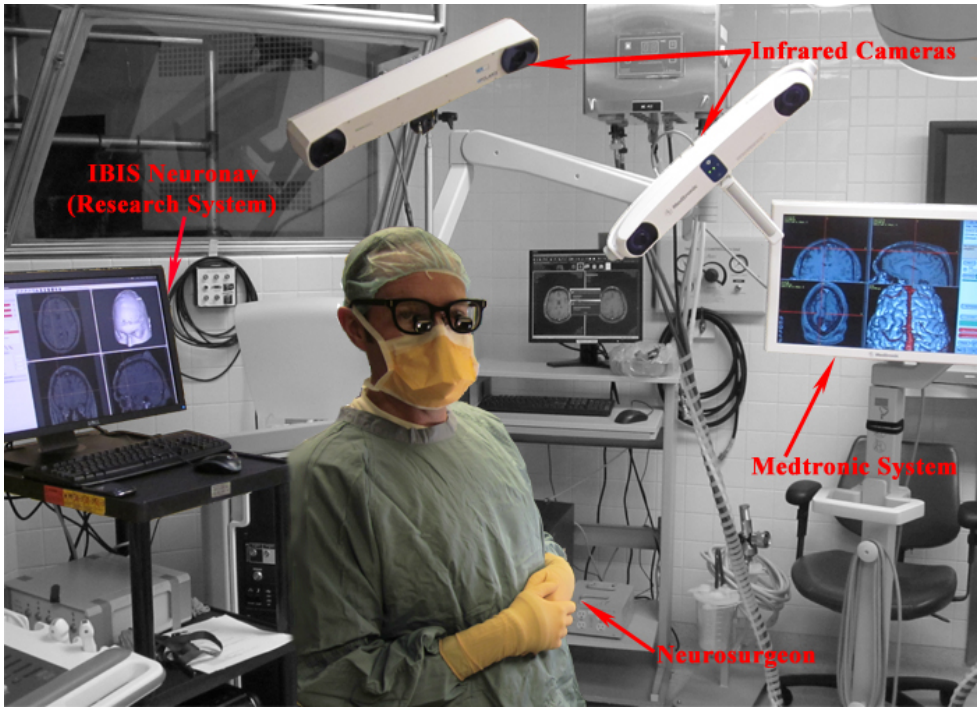

Figure 2.3 Typical setup of neuronavigation unit in the operating room of the MNH. The Medtronic (commercial) system operates in parallel to IBIS Neuronav (research system) both using infrared cameras to track reflective spheres on a reference and pointing tool (not shown).

All of these systems relate the real-world coordinates of a patient to those of the pre-operative images using a rigid body transformation. This transformation is calculated by both identifying either implanted markers or external anatomical features on both the patient and the preoperative images

and employing a least squares minimization procedure to determine the spatial transformation between the two sets of data (Comeau, et al. 2000). For many cases using optical systems, a similar result is achieved by sampling a series of points on the skin surface with the tracked pointing tool and sampling the corresponding points on the segmented 3D surface of the skin derived from the preoperative images.

2.2.1 Relating Coordinate Systems

There are several different coordinate systems involved in an IGNS intervention and they must all be related to each other in order for neuronavigation to be effectively used. This is done with a set of transformations that map the spaces of the patient and tracked tools to that of the preoperative images. Figure 2.4 illustrates the different coordinate systems generally involved in an IGNS case.

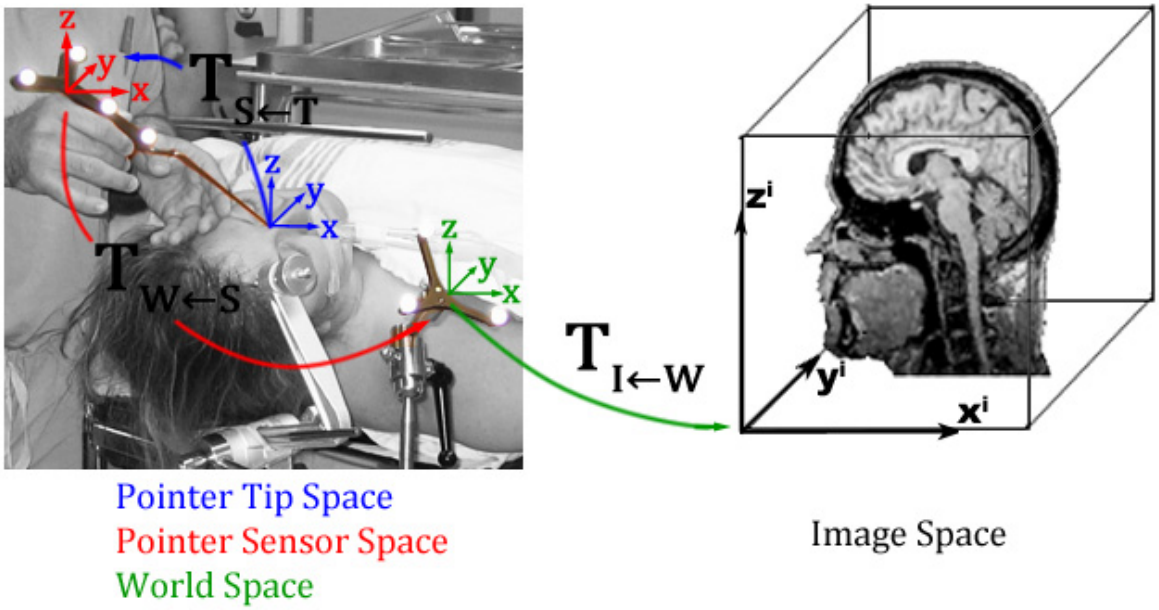

Figure 2.4 Schematic diagram showing the relationship between different coordinate systems involved in IGNS as well as the transformations that relate them to each other.

We can represent the position of the tip of the tracked pointer ($\boldsymbol{P_T}$) in our preoperative image space ($\boldsymbol{P_I}$) with the following relationship:

$$\boldsymbol{P_I} = T_{I \leftarrow W} \cdot T_{W \leftarrow S} \cdot T_{S \leftarrow T} \cdot \boldsymbol{P_T} \tag{2.1}$$

Where $T_{I\leftarrow W}$, $T_{W\leftarrow S}$, and $T_{S\leftarrow T}$ represent the rigid transformations from world space to image space, the pointer's sensors to world space, and the pointer's tip to the pointer's sensors respectfully. The $T_{S\leftarrow T}$ transformation is obtained by calibrating the tip of the pointer. This is done using a simple calibration procedure and is described in detail in the next section (Hartov, Eisner, et al. 1999) . The $T_{W\leftarrow S}$ transformation is determined by relating the position of the tracked pointer's sensors to the origin of the world (tracked) space and is obtained directly from the tracking system (Wiles, Thompson and Frantz 2004). The $T_{I\leftarrow W}$ transformation is generally referred to as the patient-to-image registration (or patient-to-image mapping). It is obtained using 7 or more landmarks on the patient's nose, eyes, and ears as well as the corresponding landmarks on the preoperative images. The anatomical landmarks are first identified on the preoperative images and then the surgeon touches the corresponding physical landmarks on the patient's skin with the pointer. Using these two sets of points $T_{I\leftarrow W}$ is computed in a least squares process between corresponding points to obtain the best mapping.

2.2.2 Tracked Pointer Calibration

To determine the relationship between the pointer's tip and its sensors ($T_{S\leftarrow T}$), a calibration procedure that involves holding the tip at a fixed location and recording its location for several different orientations is performed (Hartov, Eisner, et al. 1999). To compute the location of the tip of the pointer in tracked space, given n different positions, we have n equations of the form:

$$\boldsymbol{P_T} = M_i \times T_{S\leftarrow T} + \boldsymbol{P_{S_i}} \qquad (i = 1\ to\ n) \tag{2.2}$$

Where $\boldsymbol{P_T}$ is the location of the fixed tip point in the pointer's space, M_i the rotation matrices corresponding to the i^{th} position, and P_{S_i} the sensor positions and '×' is the cross product. The location of the pointer tip, with respect to the receiver reference frame, represents only three degrees of freedom, thus, by subtracting $(n-1)$ equations from the first equation, and knowing that $\boldsymbol{P_T}$ remains constant, then:

$$0 = (M_1 - M_i) \times T_{S\leftarrow T} + \left(\boldsymbol{P_{S_1}} - \boldsymbol{P_{S_i}}\right) \qquad (i = 2\ to\ n) \tag{2.3}$$

We can then sum the remaining $(n-1)$ equations to obtain:

$$\sum_{i=2}^{n}\left(\boldsymbol{P_{S_i}} - (n-1)\cdot \boldsymbol{P_{S_1}}\right) = \left((n-1)\cdot M_1 - \sum_{1=2}^{n} M_i\right) \times T_{S\leftarrow T} \tag{2.4}$$

Or more simply in matrix notation as:

$$P = M \times T_{S\leftarrow T} \tag{2.5}$$

Which can easily be solved for $T_{S\leftarrow T}$, thereafter allowing us to compute the location of a feature on which the pointer tip is located:

$$T_{S\leftarrow T} = (M^T \times M)^{-1} \times (M^T \times P) \tag{2.6}$$

2.2.3 Main Sources of Error

There are many different ways that error is introduced into the IGNS environment and it begins with the different hardware used. There is error associated with acquiring the preoperative images through use of MRI or CT acquisition parameters and hardware settings (Fonov, et al. 2010), (C. R. Maurer Jr., et al. 1996), the tracking system's ability to accurately and consistently report the position of a tracked tool that will influence the accuracy throughout the IGNS intervention (Khadem, et al. 2000), different tracked tools have different errors associated with their calibration and use (Hartov, Eisner, et al. 1999), there are many different types of errors associated with relating the physical space of a patient with the space in which the preoperative images are taken through registration procedures (Zitova and Flusser 2003) (van den Elsen, Pol and Viergever 1993) (Fitzpatrick, West and Maurer Jr. 1998), and finally, a plethora of errors can be introduced by biological changes in the patient throughout the surgery (Hill, et al. 1998). This section goes into some detail about some of these different types of errors and how they can affect IGNS interventions.

2.2.3.1 Pointer Calibration Errors

There is uncertainty associated in reporting the position of any tracked surgical pointer, which is introduced during imperfect calibration of these tools (Hartov, Eisner, et al. 1999). As described in the previous section, the calibration relies on using a fixed point in several different orientations. Since these points may no longer be identical, error is introduced when determining the transformation between the pointer's sensors and its tip. This error is expressed as the root mean square (RMS) of the standard deviation (σ) of the static point's x, y and z positions:

$$\epsilon T_{S\leftarrow T} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \quad (2.7)$$

Since the surgical pointer is a tool that is used throughout an entire IGNS intervention and is used to relate the physical position of a patient to the preoperative image space it is important to minimize this type of error in order to minimize its propagation through different parts of the intervention that rely on its accuracy.

2.2.3.2 Tracking Errors

There are errors associated with a tracking system's ability to accurately and consistently report the position of a tracked tool. For optical tracking systems, such as the Polaris infrared camera, one of these types of error is called "jitter", and refers to how the reported position of a static tool varies over time. The jitter varies with the type of tracking system used as well as the camera's position with respect to the dynamic reference frame (DRF) and the tracked tools (Khadem, et al. 2000). This type of error is very important during the patient-to-image registration phase of an IGNS intervention. Since the DRF remains static near the head of the patient and the surgical pointer remains static on a landmark of interest, having the lowest possible jitter will decrease the error in accurately reporting the physical location that is being targeted. There are other sources of error associated with magnetic tracking systems (Day, Murdoch and Dumas 2000), as well as ultrasonic tracking systems (O'Donnell., et al. 1994) which are related to parameters associated with their respective hardware. Optical tracking systems were used in this thesis so the errors associated with the other two systems will not be discussed in any detail.

2.2.3.3 Patient-to-Image Registration Errors

The patient-to-image registration method described earlier assumes that the patient's skin and brain behave as a rigid body; an assumption that is unfortunately false, leading to errors in registration. Registration and registration errors are complex and are categorized into different types. They are discussed in detail in section 2.3.

Since gravity affects the movement of both the patient's skin and brain, positioning of the patient in the operating room is important. Preoperative images are taken with the patient in the supine position, thus the differences due to the effects of gravity will be minimized if the patient is operated on in this same position in contrast to a prone or lateral position. Minimizing the registration errors is extremely

important since this step of the IGNS procedure determines the accuracy of the mapping between the physical location of the patient and that of their preoperative images is.

2.2.3.4 Preoperative Imaging Acquisition Errors

MRI is the most common form of preoperative imaging, so this section will focus on errors associated with acquiring with this modality only. MRI suffers from several image distortions, geometric and motion artifacts and partial volume effects; all which are dependent on the acquisition parameters. Geometric distortions arise from different sources including tissue-dependent chemical shifts and different magnetic susceptibility changes. From a hardware point of view, static field inhomogeneities can also contribute to these distortions. The largest source of geometric distortions in modern MRI system arises from the gradient field nonlinearity. Recent work has been done in order to minimize these errors (Fonov, et al. 2010). Geometric distortions can be in the order of 2-3 mm if no precautions are taken to avoid them. (C. R. Maurer Jr., et al. 1996).

2.2.3.5 Brain Shift

The major weakness of commercial neuronavigation systems is that they rely solely on preoperative images. This means that they cannot compensate for any changes that have occurred since these images were acquired. These movements and deformations in the brain tissue are referred to as ‘brain shift’ and occur due to a variety of different factors. Some of these factors include changes in intracranial pressure, changes in cerebrospinal fluid (CSF) volume, changes in brain volume, and changes in blood volume due to blood loss and different drugs used throughout the surgical intervention. Commercial IGNS systems based on preoperative data alone are not able to account for these types of deformations, which have been shown to range up to 30 mm (Hill, et al. 1998), (Roberts, et al. 1998).

2.3 Image Registration

Registration describes the process of matching two corresponding images taken at different times, from different perspectives or with different sensors (i.e. MRI and CT). Image registration is a crucial step in all image processing and analysis tasks in which the final information is gained from the combination of various data sources. In many medical contexts, registration is used to combine information from different imaging modalities to gain more complete information about a patient, monitoring tumour growth, treatment verification, and comparing patient data with anatomical atlases (Mistry and Banerjee 2012). Registration is an important aspect in IGNS interventions because it is

used to determine the relationship (mapping) between the physical location of a patient in the operating room and the corresponding location on their preoperative images in order for neuronavigation to be used. There are many different methods in which image registration can be performed; such as curve methods, correlation methods and wavelet methods, however, the most common use in IGNS is point-based rigid body registration. The discussion here is focused on this domain only.

2.3.1 Registration Framework

When performing a registration, we refer to aligning objects of interest into a common coordinate system by means of a transformation. This transformation can vary from a simple translation or scale to more complicated movements such as rotations, warping and nonlinear changes. The basic components of a registration framework include two input images, a transformation, some sort of similarity metric in which to evaluate the quality of the registration, an interpolator and an optimizer (Labadie, Davis and Fitzpatrick 2005). Figure 2.5 shows the general workflow of a registration process. We define one image as fixed and one as moving. The moving image, as its name suggests, is that which will be transformed to best match the fixed image. In the context of IGNS patient-to-image mapping, our preoperative images represent the fixed images and the patient's physical location is the moving image. Image registration is a very important aspect of IGNS interventions and there are many techniques that have been used and investigated over the past two decades (Maintz and Viergever 1998), (Zitova and Flusser 2003), (Hill, et al. 2001), (van den Elsen, Pol and Viergever 1993).

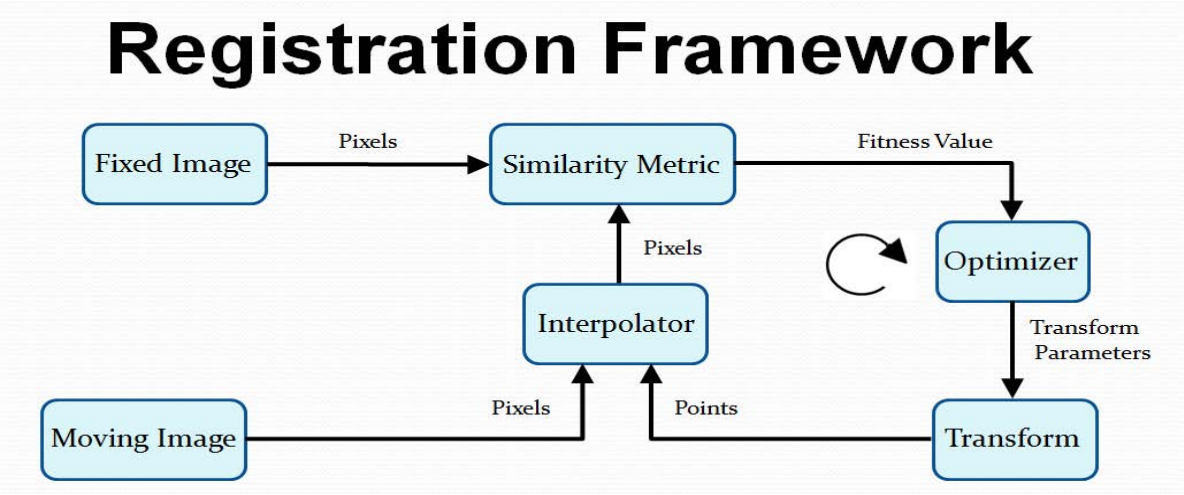

Figure 2.5 Basic components and workflow of a registration framework.

We treat registration as an optimization process trying to find the spatial mapping that will bring the moving image into the best alignment with the fixed image as indicated by a similarity metric.

2.3.2 Steps of Image Registration

There are four main steps involved in a registration process (Mistry and Banerjee 2012); the first of which is *feature detection*. Detected features range from salient to distinctive objects, such as closed boundary regions, edges, contours, line intersections, etc. and are either manually or automatically identified. These features can be represented in different ways, including their center of gravity, line endings or distinctive points within the feature. The second step is *feature matching*. This refers to the correspondence between features in the two images, and is a large area of research for matching images of different modalities. Features can be matched in different ways; by means of the image intensity values in their close neighborhoods or by the feature spatial distributions, for example. Once features have been detected and matched, the *transformation model estimation* step is performed. This involves estimated the parameters of the mapping function to align the moving image with the fixed image based on the features that were detected in the first two steps. In general, there are two main types of mapping functions:

i) Global mapping: using all feature points to estimate one transformation that will be applied to the entire image.

ii) Local mapping: treating the image as a composition of blocks and estimating transformations for each of the individual blocks.

The final step of the registration process is *transformation application and image resampling*. The moving image is transformed by means of the mapping function estimated in the previous step. Since the transformation can cause image values to be in non-integer coordinates, these values are computed by an appropriate interpolation technique such as nearest neighbor, bilinear, bicubic, quadratic splines, Gaussian or higher order B-splines.

2.4 Registration Errors

One of the major limitations IGNS is the inaccuracy of the registration between an image and the physical anatomy of a patient. Point-based rigid-body registration is the most frequently used procedure in IGS and its goal is to find an optimum transformation to map two sets of points localized

in a patient before and during an operation. The accuracy of a registration procedure is assessed by measuring the length, such as the root mean square (RMS), of a point-pair vector between image and physical anatomy target locations. Recent literature suggests three main error metrics when analyzing the accuracy of point-based registration methods: i) fiducial localization error (FLE) – the error in locating fiducial points, ii) fiducial registration error (FRE) – the distance between corresponding fiducial points after registration, and iii) target registration error (TRE) – the distance between corresponding points other than the fiducial points after registration (Moghari and Abolmaesumi 2010), (Labadie, Davis and Fitzpatrick 2005), (Liu, et al. 2009). Fiducials are important as they are the basis of registration in IGS systems and are important for measuring errors. Fiducials come in several different forms including, bone-implanted markers, proprietary head frames, skin-affixed markers, anatomical landmarks, and skin surface contours. Figure 2.6 shows an example of these various types of registration errors and they are described in more detail in the following sections.

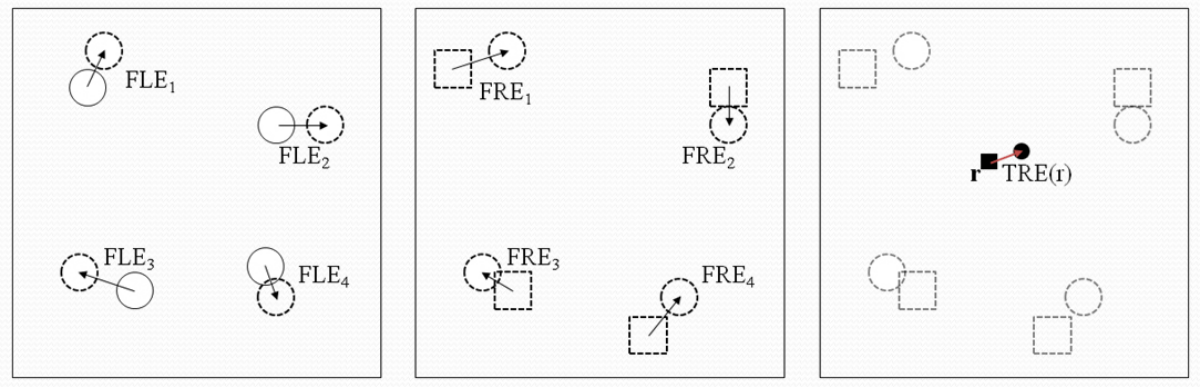

Figure 2.6 Various types of registration error. The FLE measured at each fiducial is the distance between the true position (solid circle) and the measured position (dashed circles) of the fiducial. The FRE measured at each fiducial is the distance between the measured position of a fiducial in one space and its counterpart (dashed circle and dashed square), after registration. The TRE, measured at a point r relative to some given origin, is the distance after registration between the anatomical location (solid square) represented by r in one space and the corresponding anatomical point in the other space (solid circle). [Figure reproduced from (Fitzpatrick, West and Maurer Jr. 1998)]

2.4.1 Fiducial Localization Error (FLE)

Registration of preoperative images to intraoperative surgical anatomy requires identification of landmarks both in the image and on the patient in the operating room. It is impossible to accomplish this perfectly due to errors that occur when locating each fiducial marker. The fiducial localization

process can be based on interactive visual identification of anatomical landmarks, such as the junction of two linear structures, e.g., the central sulcus with the midline of the brain. Alternatively these features can be markers attached to the anatomy, designed to be detected accurately by means of automatic algorithms. Regardless of which method is used, the chosen point will inevitably be erroneously or imperfectly chosen. Different anatomical landmarks are associated with different FLE distributions spanning the range of 0.5 ± 0.5 mm to 2.0 ± 2.1 mm on an MRI image depending on different image characteristics (i.e. voxel size), the type of sequence used and the person choosing the landmarks, and in the range of 1.9 ± 1.0 mm to 3.2 ± 1.6 mm in the physical domain depending on the surgeon choosing the point, and the type of landmark used (Liu, et al. 2009). FLE can be caused in both the image domain (I-FLE) and the physical domain (P-FLE). One can estimate the FLE by calculating the average of the measured distance value between repeated localizations of one fiducial point. Using I-FLE as an example:

$$\mathrm{I}-FLE = \frac{1}{n-1}\sum_{i=1}^{n-1}\left\|f_{i+1}^{I} - f_{i}^{I}\right\| \tag{2.8}$$

Where f_i^I is the image coordinate of the fiducial point and n is the number of times the fiducial is repeatedly chosen. There are many factors that contribute to the value of I-FLE being nonzero including image distortion, signal-to-noise ratio (SNR), resolution, human error and potential errors in computer algorithm used to detect landmark fiducials. Within the operating room, infrared or electromagnetically tracked tools are used to determine the location of a tool placed on a fiducial marker. Factors that contribute to the error in P-FLE include human error in placement of the tool on the patient and error associated with the tracking system. These types of errors are difficult to observe directly, but they can be observed indirectly through the registration errors that they contribute to.

2.4.2 Fiducial Registration Error (FRE)

Once the fiducial markers have been localized on both the image and on the patient they are used to register the two images. Due to FLE, the exact location of the fiducials is not known, resulting in an imperfect registration. Mathematically, a rigid transformation matrix T^{W-I} is calculated to minimize the RMS distance between the landmark pairs:

$$T^{W-I} = \arg\min_{\mathrm{T}} \left(\sqrt{\frac{1}{N}\sum_{i=1}^{N} \| T(f^{w}) - f^{I} \|^{2}} \right) \tag{2.9}$$

Where f^W is the fiducial point in world coordinates and N is the number of fiducial features used in the registration. Any nonzero displacement between a transformed fiducial point and its corresponding fiducial point in the target space is an FRE. A common measure of overall fiducial misalignment is the RMS error (West, et al. 2001). We can describe this mathematically as:

$$FRE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \omega_i^2 \left\| T^{W-I}\left(f_i^I\right) - f_i^W \right\|^2} \tag{2.10}$$

Where ω_i^2 is a non-negative weighting factor that can be adjusted to decrease or increase the influence of different fiducials. This weighting factor can be used to decrease the influence of less reliable fiducials. To illustrate, we can set $\omega_i^2 = {}^1/_{FLE_i^2}$, where $\mathrm{FLE_i}$ is the fiducial localization error for fiducial i. This would result in having fiducials whose localization error is smaller, and thus prone to less error and more reliable, have a greater influence and the converse for those points whose localization is more difficult and less reliable. The term $\left\| T^{W-I}\left(f_i^I\right) - f_i^W \right\|^2$ can be considered the individual fiducial registration error for the i^{th} fiducial.

It is important to note that all of the anatomical points are aligned based on the localization of the fiducials. For a given system, if the fiducial fit is consistently poor, the alignment of the other points is likely to be poor. There are situations, however, where using a large number of fiducials that are well spread through the domain of interest that can yield good alignment for the rest of the anatomy, leading to a lower TRE, if the error for each fiducial is similar. The converse of this holds true as well; a consistently good fiducial fit is likely to yield good alignment of other points. Another important point: a small FRE for a specific case does not translate to more accuracy for that case. FRE is used as a general gauge of good registration and there is usually a threshold for which the FRE must be to ensure an acceptable level of accuracy for surgery. Having FRE dramatically below this threshold does not indicate that more accuracy is achieved (Labadie, Davis and Fitzpatrick 2005). For example, if an arrangement of fiducials was chosen completely on one side of the domain of interest there may be very good alignment on this side for these points, however, lever effects could cause a

larger shift on the opposite side of the domain leading to a high TRE. A final point of interest is in regards to the number of markers used for the registration. One would expect an increase in accuracy with an increased number of markers. Counter to intuition, an increased number of markers usually results in larger FRE since it becomes more difficult to perfectly align multiple markers, however, once a sufficiently large number of markers have been used the mean FRE will plateau to a steady value (Fitzpatrick, West and Maurer Jr. 1998).

2.4.3 Target Registration Error (TRE)

Once the best fit transformation has been determined and the fiducial registration completed, tracking begins and the locations of different targets of interest can be compared by identifying the anatomy on the patient and its corresponding location on the preoperative image. Any difference between the location of the physical anatomy and the point on the image is the TRE. Similar to FRE, TRE depends on the FLE, however, TRE is not the same for all points within the image and it can vary depending on the position of the point of interest relative to the fiducial markers and the configuration of the markers. Mathematically we can describe TRE by the following:

$$TRE(t^I) = \|T^{W-I}(t^I) - t^W\| \tag{2.11}$$

Where t^I and t^W correspond to the target point positions in the image coordinates and world coordinates respectively. To minimize TRE, strategic placement of fiducials is important. The distribution of the TRE has been reported to improve when avoiding collinear and coplanar placement of markers, using as many markers as possible, keeping markers as far apart as possible and placing markers so that they surround the surgical target (West, et al. 2001). It is important to note that TRE is a statistical measure. Since the registration is a best overall fit to fiducial markers whose localization error is random, the TRE is random as well. Thus, TRE is reported as a statistical distribution: TRE is χ^2 distributed with 3 degrees of freedom (Fitzpatrick and West 2001). This means that the TRE for a target point falls within this distribution and ninety-five percent of the time, TRE can be expected to be less than approximately 1.7 times its mean value. However, outliers may exist. To illustrate this, refer to Figure 2.7, which shows a χ^2 distribution with a mean of 1. This is analogous to a system having a reported TRE of 1 mm; the accuracy of the system is on average 1 mm, but it may have considerably larger or smaller values. The TRE is the most important error as it represents the error between an anatomical location and its homolog in image coordinates

transformed to the patient space. TRE can be viewed as a vehicle trying to reach a specific destination. The distance between where the GPS says he is and his actual location is the TRE.

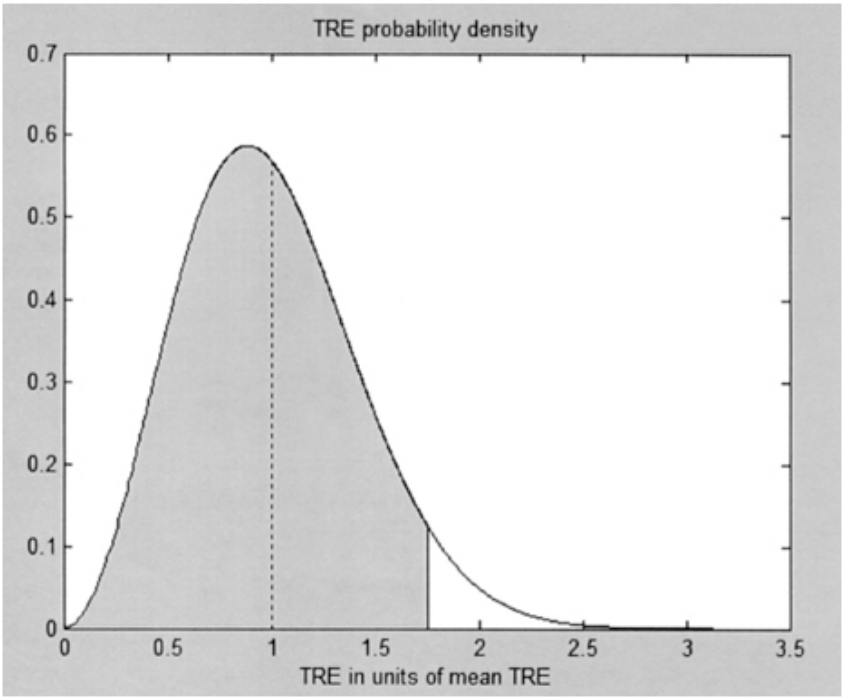

Figure 2.7 The horizontal axis is TRE in units of mean TRE with the vertical line at the mean. The shaded area represents 95% area with the right edge at 1.75. Thus, 95% of the time TRE can be expected to be less than 1.75 times its mean.[Reproduced from (Labadie, Davis and Fitzpatrick 2005)]

2.4.4 An Illustrative Analogy

An analogy to understand FLE, FRE, and TRE can be made to fitting a knight with a suit of armor (Labadie, Davis and Fitzpatrick 2005). The fiducial markers in this analogy would be key points needed to fit the suit to the knight such as neck size, arm length, waist size, inseam, and chest size. The FLE can be separated into the FLE of the suit – accuracy of the construction of the armor such that the sizes are not exact, and FLE of the knight – accuracy of the measurement of the sizes. Putting the suit on would be analogous of registration and the fit at each of these fiducial marks being FRE. TRE would correspond to the fit of the suit in places that were not measured, such as around the thigh or forearm. The suit will fit better in some areas than in others, and if the suit making process was

repeated with random measurement errors, some suits will fit better than others; thus a distribution, both spatial and statistical, describes the fit better than a single number.

2.4.5 Relationships Between FLE, FRE, and TRE

The properties of FRE and TRE and their relation with FLE have been under investigation for several decades. In 1979, Sibson was the first to demonstrate that there was a relationship between these three error measurements (Sibson 1979). He showed that the relationships among the expected values (denoted hereafter by $\langle\cdot\rangle$) of FLE^2, FRE^2, and TRE^2 are known to an excellent approximation for the most common case, in which the individual FLE are random, their probability distributions are isotropic, independent and identical for all N fiducials, and a uniform weighting ($\omega_i^2 = 1$). The simplest relationship is that between $\langle FLE^2\rangle$ and $\langle FRE^2\rangle$:

$$\langle FRE^2\rangle \approx \left(1 - \frac{2}{N}\right)\langle FLE^2\rangle \tag{2.12}$$

An important implication of equation 2.12 is that FRE is completely independent of the fiducial configuration. For a given number of fiducials and a given distribution of FLE, FRE^2 will be distributed in the same way regardless of whether the fiducials are clustered within a few centimeters of each other or are kilometers apart. It also means that planar, linear or regular polyhedral arrangements of fiducials exhibit no differences with regard to FRE. As mentioned earlier, in most medical applications the TRE is of more concern and unlike $\langle FRE^2\rangle$, $\langle TRE^2\rangle$ depends on the fiducial configuration and also on the target positions as well. As described by Fitzpatrick *et al.* in 1998, the relationship between $\langle TRE^2\rangle$ and $\langle FLE^2\rangle$ is most easily stated in terms of quantities measured relative to the principal axes of the fiducial configuration:

$$\langle TRE^2\rangle \approx \left(\frac{1}{N}\right)\left(1 + \frac{1}{3}\sum_{k=1}^{3}\frac{d_k^2}{f_k^2}\right)\langle FLE^2\rangle \tag{2.13}$$

Where d_k is the distance of the target from principal axis k, and f_k is the RMS distance of the fiducials from the same axis (West, et al. 2001). From the relationships in (2.12) and (2.13), we can see that both $\langle TRE^2\rangle$ and $\langle FRE^2\rangle$ are proportional to $\langle FLE^2\rangle$, however, $\langle TRE^2\rangle$'s dependence on fiducial configuration makes it a much more powerful and superior figure of merit for registration accuracy. To illustrate this, consider an approximately linear fiducial configuration. This will result in an extremely small value of f_k in (2.13), so, a target marker at an appreciable distance from this line

of fiducials (corresponding to d_k) will be large resulting in a large $\langle TRE^2 \rangle$ even if $\langle FLE^2 \rangle$ is small. For this same configuration, however, $\langle FRE^2 \rangle$ will be small. In contrast, if the target is located near the centroid of a non-linear fiducial configuration, then all the d_k will be small, none of the f_k will be small, and $\langle TRE^2 \rangle$ can be smaller than $\langle FRE^2 \rangle$ depending on $\langle FLE^2 \rangle$.

Counter to intuition, the registration error FRE_i for a specific marker, i, is a perverse indicator of TRE in the vicinity of that marker (West, et al. 2001), as shown by the following relationship:

$$\langle TRE^2\left(t_i^I\right) \rangle \approx \langle FLE^2 \rangle - \langle FRE_i^2 \rangle \tag{2.14}$$

Consequently, for a given configuration of fiducials, TRE tends to be worse in the vicinity of those fiducials whose alignments tend to be better. This happens because the largest motion of a fiducial occurs for those that lie furthest from the centroid of the configuration. For example, a small change in the form of rotation about the centroid will have little effect on the fiducials located close to the centroid but the relative large motion at a point distant from the centroid may bring a fiducial pair into closer proximity, reducing FRE_i. While the overall goodness-of-fit is expected to improve by this rotation due to a single fiducial pair, the error from the rotational misregistration becomes large as the distance from the fiducial centroid increases. It is important to note that equations (2.12), (2.13) and (2.14) ignore any terms involving FLE^4 and higher and that they apply both to the situation in which $\langle FLE^2 \rangle$ arises from the distribution of FLE in one space with the FLE in the other being negligible and also to the situation where the error is in more than one space, in which $\langle FLE^2 \rangle$ is simply the sum of the $\langle FLE_j^2 \rangle$ of the j spaces (Fitzpatrick, West and Maurer Jr., Predicting Error in Rigid-Body Point-Based Registration 1998).

Over the past thirty years there have been many studies that have tried to determine a relationship between TRE and FRE. It has been common practice for many years that FRE for a given procedure within an image guidance system can be used as an indicator of the expected level of accuracy for the rest of procedure. However, in 2009, Fitzpatrick demonstrated that this assumption is wrong (Fitzpatrick 2009). Fitzpatrick proved that variations in FRE and TRE are uncorrelated and to first order are completely independent of one another. This result, however, does not remove any of the importance from equations (2.12) and (2.13) which relate expected values. Equation (2.12) can be used to estimate the FLE distribution from measurements of FRE which can in turn be used to

estimate TRE for a given case, but the FRE for the given case plays *no* role in the estimate of TRE for that case (Fitzpatrick 2009). Despite this, there is still worthwhile information that can be extracted from an FRE measurement. A larger than normal FRE for a given case indicates a larger than normal probability that the registration system has been compromised for that case. Based on work in (Fitzpatrick and West 2001), a probability can be formalized to support whether or not a system is working properly. Thus, if an FRE that is slightly smaller, or slightly smaller than usual it is a good indication that the system is functioning properly and equation (2.13) can be used to predict registration accuracy but that value of FRE gives no additional information whatsoever about the actual registration accuracy for that case.

2.5 Previous work in IGNS error analysis

IGNS and neuronavigation systems have been around for many years now and much work has been done to attempt to characterize some of the errors and limitations associated with these types of systems and interventions. The main point of using a neuronavigation system during an IGNS intervention is to improve the accuracy of the surgery. Inaccuracies of a few millimeters can make the difference between a successful and a less successful surgery (Stieglitz, et al. 2013). As described earlier, neuronavigation relies on registration between a patient's physical anatomy and the preoperative images. Many different types of neuronavigation systems are available and their accuracies have been reported to vary between 0.7 mm and 5 mm, in terms of patient-image registration at the beginning of surgery. Table 2.1 summarizes some of the different research groups that use neuronavigation systems and point based registration methods, along with their reported errors. It is easily seen that when surgically implanted screws are used, the registration error is at its lowest, however, this is a very invasive procedure and is not feasible for every type of procedure due to time, cost, and other factors. When landmarks or fiducials are used, the registration accuracy is generally not better than 2.5 mm and for procedures involving eloquent areas of the brain, exceptional accuracy is required and these current systems are insufficient at providing this accuracy forcing a neurosurgeon to rely on his mental map of the patient's neuroanatomy, indicating a need for improvement in this field if neuronavigation is to become a more reliable tool for surgery. In addition, it has been shown that accuracy decreases as the length of surgery increases due to a number of surgical, biological and technical factors related to the patient, the tools used, and environmental changes (Stieglitz, et al. 2013), (Golfinos, et al. 1995), (Germano, et al. 1999).

Table 2.1 Patient-to-Image registration accuracy in the Literature

Reference	System	Paired Point Matching (mm)		
		Landmarks	Fiducials	Screws
(Mercier, Del Maestro, et al. 2011)	IBIS Neuronav	4.9 ± 1.1		
(Sipos, et al. 1996)	FARO Surgicom	2.7		
(Pfisterer, et al. 2008)	Stealth Station (Medtronic)	4.0 ± 1.7		
(Sipos, et al. 1996)	FARO Surgicom		2.8	
(Pfisterer, et al. 2008)	Stealth Station (Medtronic)		3.5 ± 1.1	
(Gumprecht, Widenka and Lumenta 2002)	BrainLab VectorVision		4 ± 1.4	
(Pillai, Sammet and Ammirati 2008)	Stryker Navigation			0.91 ± 0.28
(Thompson, et al. 2011)	Stealth Station (Medtronic)			1.3 ± 0.5
(Brinker, Arango and Kaminsky 1998)	Zeiss MKM			0.7 ± 0.2

Chapter 3

Methods

The goal of this thesis is to quantify and characterize tracking accuracy in image guided surgery using the IBIS neuronavigation system (Mercier, Del Maestro, et al. 2011). In this chapter, details regarding the different materials used and created, as well as the experiments that were performed to quantify and characterize different errors that arise in the context of IGNS will be described. The first section describes the materials used for tracking and neuronavigation as well as the construction of a linear testing apparatus (LTA) that was used to make most of the error measurements.

3.1 Materials

Tracking was performed using one of two Polaris infrared optical cameras (Northern Digital, Waterloo, Ontario, Canada). The camera used for experiments is indicated as ‘C1’ or ‘C2’. There were up to 5 different tools passively tracked throughout all of the experiments; a TA002-135 *Pass Trax Tracker* dynamic reference frame , an NDI passive pointer and a Traxtal FS316 passive pointer, a Traxtal TA001-133 marker tool, and a Traxtal TA201-4005 marker tool (Traxtal Technologies Inc. Toronto, Ontario, Canada) . All tools had reflective spheres rigidly fixed to them. Figure 3.1 shows the different tools used to perform the experiments in this thesis.

There is a range of positions inside of which the tracking apparatus is completely visible to the camera and where tracking measurements can be made. This region is called the digitizing volume. The C1 and C2 have a cylindrical digitizing volume (see Fig. 3.2). The front of the cylindrical volume is a sphere of radius 0.5 m, centred 1.9 m from the position sensor. The back of the measurement volume is a cylinder of radius 0.5 m extending from 1.9 m to 2.4 m from the position sensor. This is illustrated in Figure 3.2 highlighting the camera’s field of view and the calibrated volume. When

tracking moving tools, they must remain in the camera's measurement volume and preferably away from the back edges of this volume which tend to be less accurate, as indicated by the manufacturer. The software used with the tracking hardware was a prototype neuronavigation system developed locally in the Image Processing Laboratory of the McConnell Brain Imaging Center at the Montreal Neurological Institute, called the Intraoperative Brain Imaging System (IBIS) (Mercier, Del Maestro, et al. 2011).

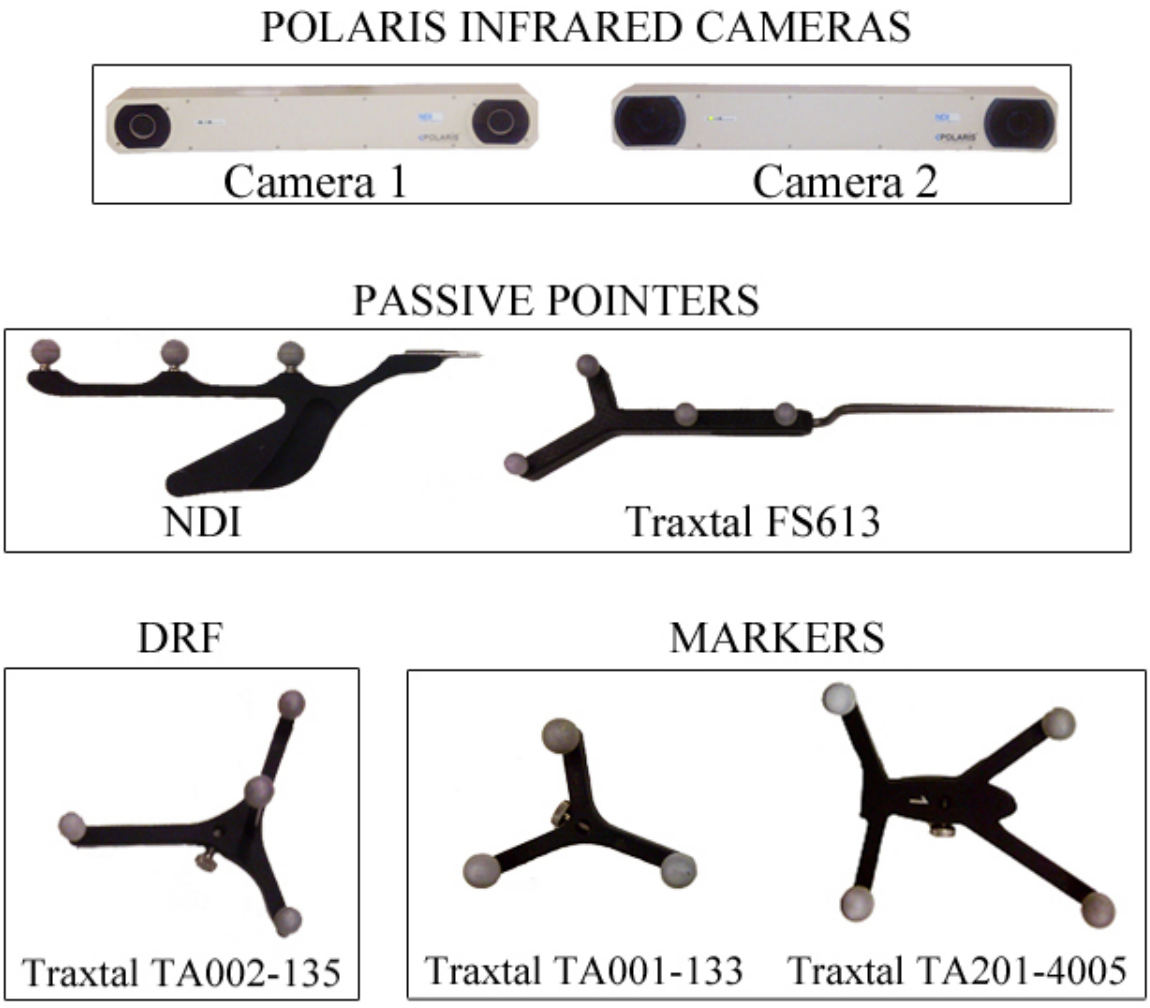

Figure 3.1 All the trackers and tracked tools used for measurements in the described experiments.

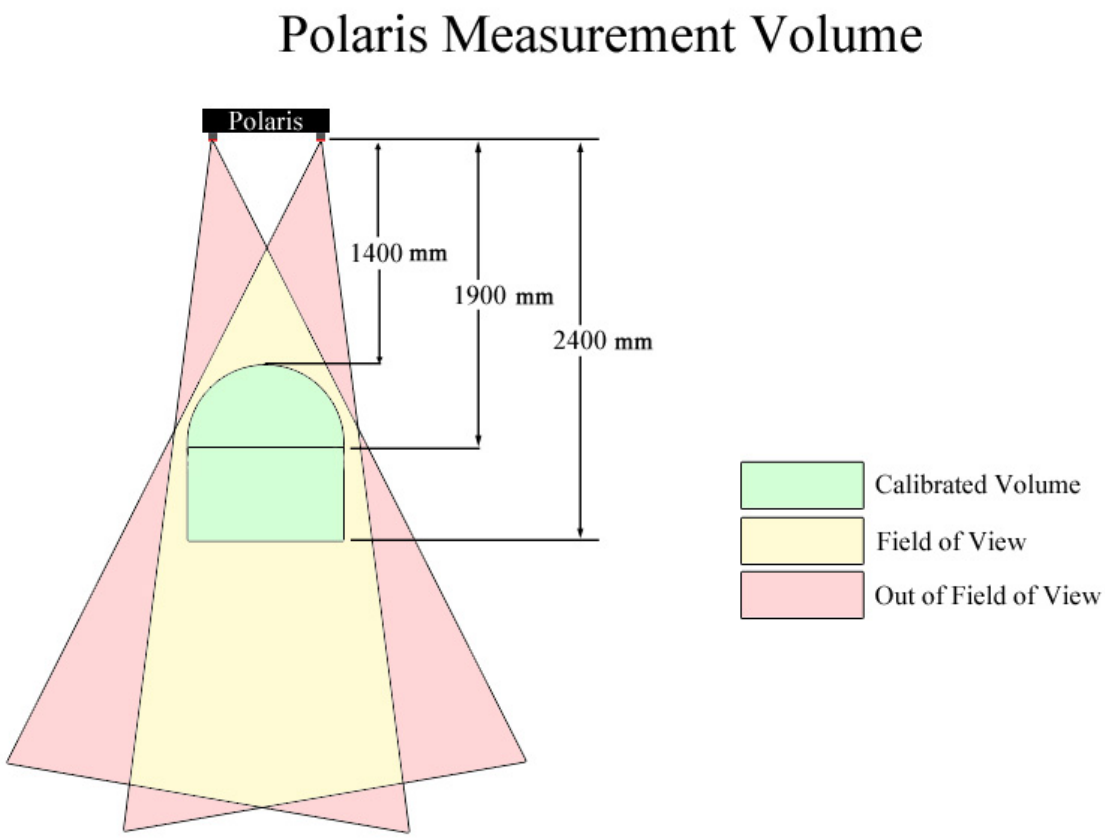

Figure 3.2 A schematic diagram showing the cylindrical measurement volume of the Polaris cameras. Green represents the calibrated volume of the camera where measurements are to be taken; yellow is the field of view of the camera, defined by the overlapping of the field of view of each of the two individual cameras and red representing the volume that is out of view. This information was provided by the manufacturer.

3.1.1 Linear Testing Apparatus

The volume and registration measurements in the following experiments were obtained through the use of a precision-machined linear testing apparatus (LTA), shown in Figure 3.3. The design was based on the work of Khadem in a paper in 2000 analyzing tracking error of different tracking systems (Khadem, et al. 2000). Due to some equipment limitations the design varied slightly and is described in the following paragraph. The purpose of the LTA was to create a volume that could be used with tracking equipment and tools and whose manufacturing errors were far less than the error expected from the actual tracking error measurements. It was also made to mimic a typical surgical

volume in a real IGNS case so that tracked tools, such as the NDI and Traxtal pointers, could be used to register the volume with a corresponding virtual image volume to observe and measure potential registration errors.

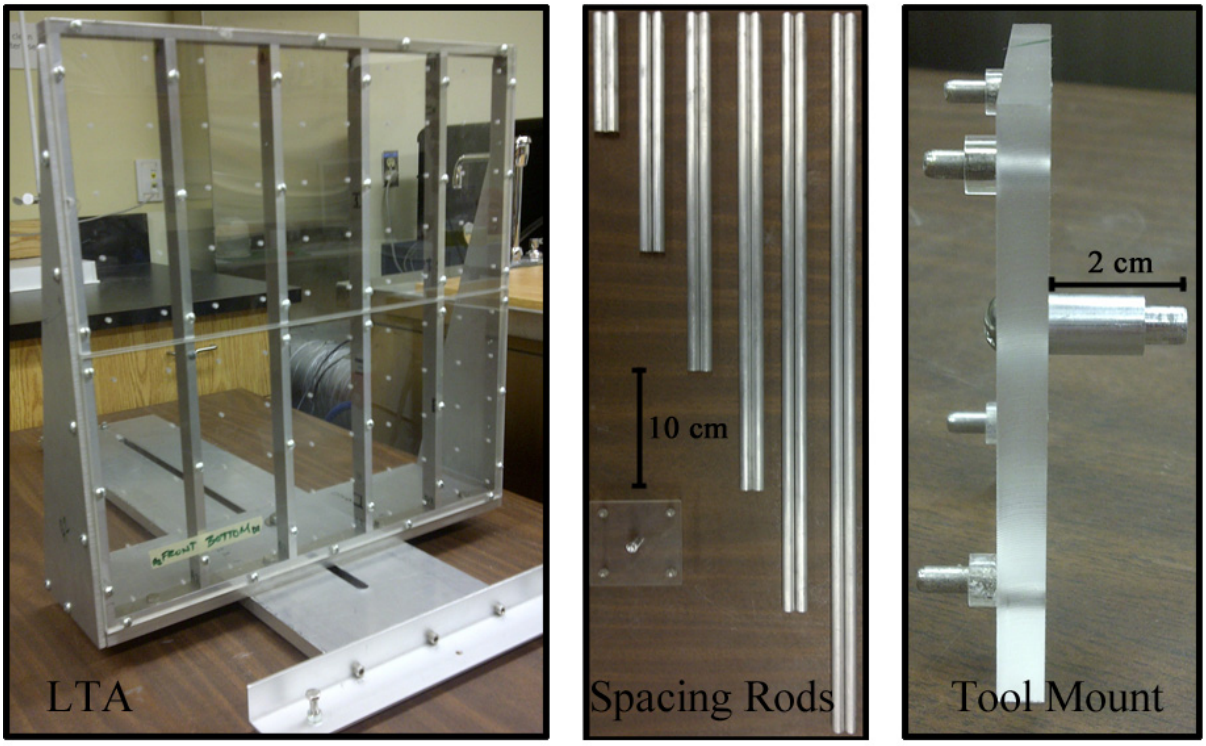

Figure 3.3 Left: view of the LTA. Middle: precision milled spacing rods ranging from 10 cm to 60 cm in length. Right: Close up side view of the mount for DRF or marker tool that mate with the holes of the LTA.

The LTA consists of a 500 mm x 500 mm target plate with a 10 x 10 grid of uniformly spaced holes. The holes are separated by 50 mm in both directions and were milled with a 0.005 mm tolerance. The holes were milled so that the Traxtal and NDI pointers could fit snugly inside without moving to enable recording of the tracked position of each hole. The target plate was made from ½" plexiglass because of the material's ability to be easily milled with the equipment available and this material is malleable enough so it would not break when ensuring that the plate was straight. Due to machine equipment restraints, the target plate was constructed from 2 identical pieces that were milled simultaneously and then put together using a third piece to ensure proper alignment (See Figure 3.4). The target plate was mounted onto a ¾" aluminum frame and side panels and screwed together on a

framing table to ensure that the target plate remained perpendicular to a flat surface in all areas. This can be seen in the left picture of Figure 3.3. Aluminum was chosen since it was a harder and heavier material than the plexiglass and would be a good anchor to attach to the plate as a frame and the additional weight would help ensure that the plate would not move when performing experiments. An additional mount with pegs that mate with the holes was made so that the DRF or marker tools could be mounted on the plate for certain experiments (far right picture on Figure 3.3). The mounting tool was 6 cm x 6 cm with a rod that extends normal to the tool, which is 2 cm in length, that mates with the tracked tools. The mounting tool with pegs was also used as a tool to verify that all holes were separated by the proper distance on the entire plane of the target plate. Since it was also milled with a 0.005 mm tolerance, the four holes with pegs that fit tightly into them were placed in each set of holes in the target plate to ensure a snug and accurate fit along the entire area.

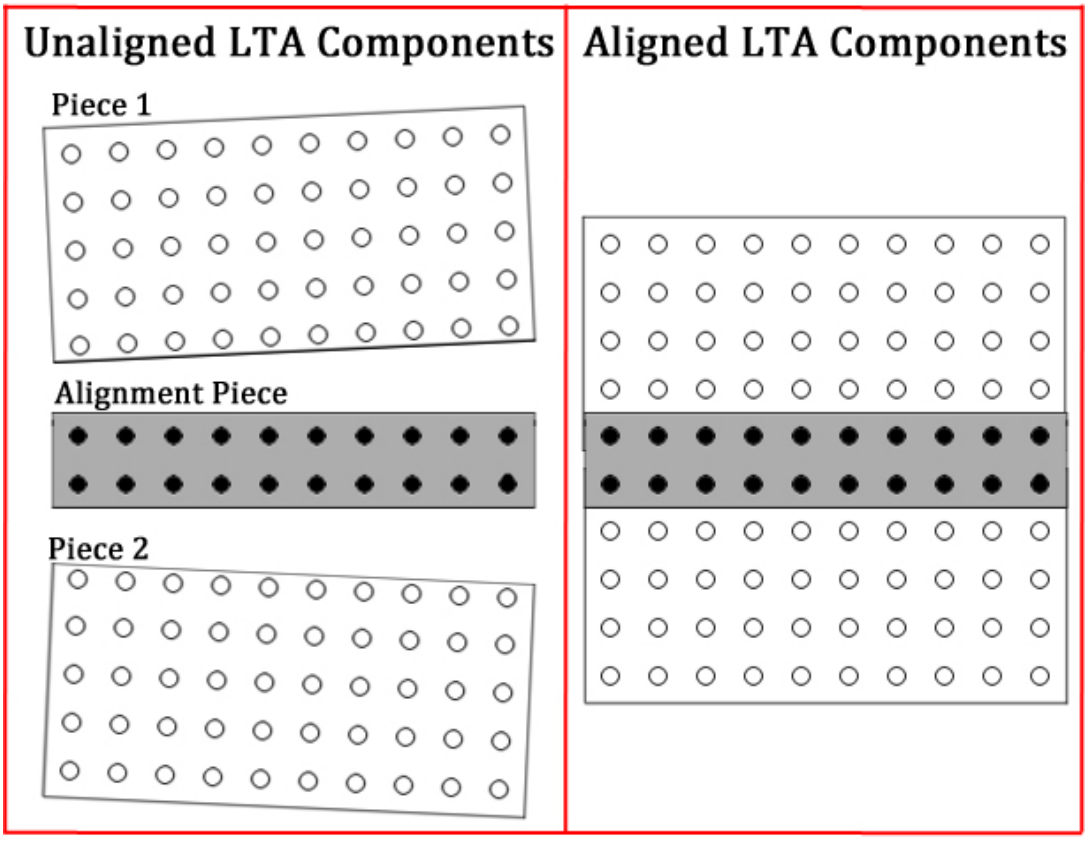

Figure 3.4 Left: 2 identical pieces of the LTA with the alignment piece before assembly. Right: LTA assembled with alignment piece ensuring proper separation between the two pieces. The open holes are represented by white filled circles and the pegs by black filled circles. The alignment piece is only used to assemble the LTA and does not remain in the finished LTA.

In order to obtain measurements in all 3 dimensions, the plate is also mounted on an aluminum track that allows movement in the direction normal to the plane of the target plate. The distance is determined through use of precision milled spacing rods that go at the edge of the LTA track and push up tightly against the bottom portion of the frame to ensure the plate remains straight when changing distances. Six pairs of rods were milled, the first measuring 10 cm and each increasing in length by 10 cm with a tolerance of 0.01 mm. Each pair of rods was milled simultaneously to avoid any additional manufacturing errors. Their precision was tested by comparing the length reading on both the lathe that was used to manufacture them as well as a set of Vernier calipers. They can be seen in the middle picture of Figure 3.3. The front edge of the track was also constructed with a lip so that the DRF or other tools could be attached to it for different experiments or to use as a non-measurement portion of the apparatus that could be used to clamp it down to a flat surface to ensure no movement of the LTA was occurring during measurements.

3.2 Pointer Calibration Error

For these experiments, the LTA was placed on a flat surface with the camera in front of it and the target grid parallel to the *x-y* plane (z being the distance from the camera) of the camera. Different properties of the error associated with calibrating the pointer were investigated through the use of different tracked pointers having different reflective sphere configurations, as well as different camera setups. The Traxtal pointer has four reflective spheres in a non-coplanar arrangement.The NDI passive pointer has three reflective spheres aligned collinearly. They can both be seen in Figure 3.1 as well.

The first experiment intended to characterize the pointer calibration error as a function of distance from the tracking camera. The DRF was placed in the centre of the digitizing volume of the camera which was positioned at the side of the LTA (See Figure 3.5). Both pointer tools were then calibrated 10 times. The plate was then shifted 10 cm along the z-axis, away from the camera using the precision milled rods, and the pointer tools were calibrated another 10 times. This was repeated for intervals of 10 cm for a range of 60 cm from the front of the calibration volume of the camera.

The second experiment intended to characterize the pointer calibration error as a function of distance from the DRF. This was performed by fixing the DRF to a non moving portion of the LTA at the front of the calibrated volume of the camera and doing the same set of measurements as described above. All of these measurements were repeated with the second camera as well.

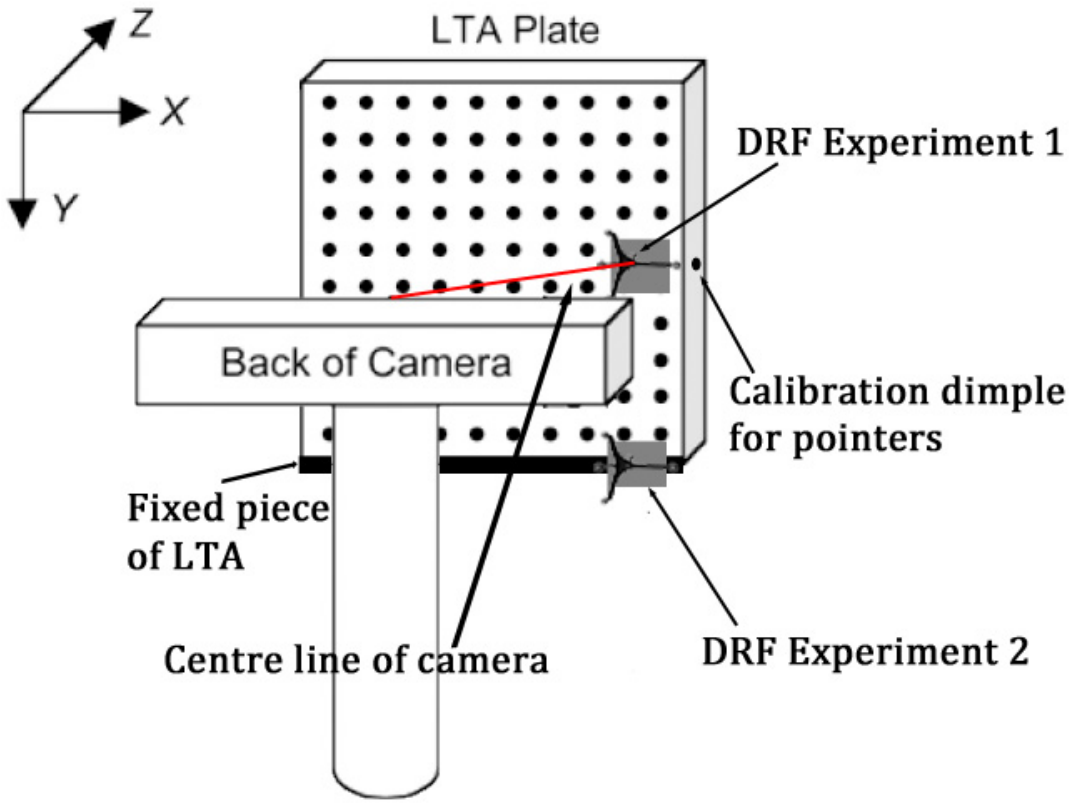

Figure 3.5 Schematic diagram of the setup for pointer calibration experiment one and two.

3.3 Tracking Error Quantification

For these experiments, the LTA was placed on a flat surface in front of the camera on a tripod with its plane parallel to the *x-y* plane of the camera. The DRF was clamped to the flat surface that supported the LTA. A laser pointer, attached to the centre line of sight of the camera, was positioned four times: each time pointing to one of the four corners (See Figure 3.6). In order to build the entire measurement volume, the camera's centre line was moved to one of four positions relative to the LTA: The first hole of the first row, the sixth hole of the first row, the first hole of the sixth row and the sixth hole of the sixth row. This can be seen in Figure 3.6 (right) as the top left corner of each of the squares labeled A,B,C and D. The third dimension of the volume was sampled by changing the distance of the LTA target plate from the front of the camera by the use of precision milled rods to extend through the entire calibrated volume of the camera. It is assumed that the digitizing volume is symmetric about the x and y axes, therefore measurements only need to be recorded in a single quadrant (Khadem, et al. 2000).

Once the camera is in the proper parallel position, 100 data points were taken with a marker tool, tracked by the Polaris camera, at the first position. Next, the marker tool was moved 5 cm (one hole) to the right to collect another 100 data points. This was repeated for 5 positions on 5 different rows, recording from a total of 25 different locations. This is equivalent to the first portion of the section labeled A in Figure 3.6 (right). The plate was then shifted 10 cm along the z-axis, away from the camera using the precision milled rods, and the entire process was repeated for each new depth through the entire calibrated volume of the camera. Once all the depths for position A had been sampled the camera would be repositioned so that sections B, C and D could be collected in the same way.

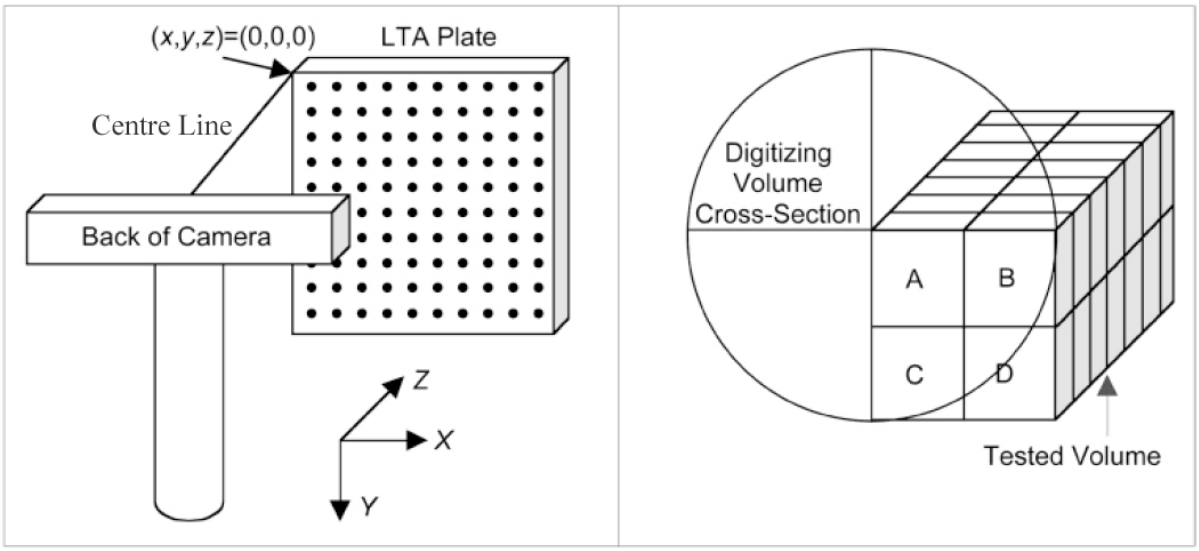

Figure 3.6 Left: The relative position of the camera and the LTA plate along with a coordinate system. Right: The actual digitizing volume of the camera. [Reproduced from (Khadem, et al. 2000)]

To quantify the tracking error associated with the camera, a parameter called "jitter" (Khadem, et al. 2000) was used. The jitter, J, is represented by the standard deviation of the Euclidean distances between the actual point location and the mean recorded location of the 100 data. The components J_x, J_y and J_z are the normalized standard deviation of each of the *x*, *y*, and *z* coordinates. These are estimates of the total jitter and its components along each axis for a given location (Khadem, et al. 2000).

This experiment was performed using two different cameras (described earlier) as well as two different marker tools; one with 3 spheres and another with 4 spheres. These can all be seen in Figure 3.1.

3.4 Error Distribution in the LTA Volume

The LTA was designed to mimic the size of a volume that would generally be tracked and operated on during an IGNS intervention. However, due to the different equipment used throughout the surgical interventions, the position of the camera, relative to the patient, varies significantly from case to case. The following experiments were performed to obtain a distribution of the error within the surgical volume for different camera arrangements. For all of these measurements, the LTA was placed on a flat surface. The DRF was placed at the front of the track of the LTA, since this is where the DRF would be placed during an IGNS intervention.

Ten different camera positions were used. The base of the tripod was placed 1400 mm from the flat surface where the LTA was located. Five different locations each with two different heights were used and chosen so that the entire volume of the LTA could be sampled without changing the position of the camera. The positions are summarized in Table 3.1 along with a schematic diagram in Figure 3.7.

After positioning the camera, 100 points at 6 different depths in the LTA were sampled using the Traxtal pointer and then registered to a corresponding volume of points in a virtual image. This volume was also registered using only the outer border of points of the LTA volume and finally with the outer border points of the LTA plus two central points of each plane. The camera position was then varied into the next position and all measurements were repeated. These measurements were also repeated for both cameras. For each set of measurements the best fit transformation was determined between sampled points and the corresponding grid volume points by means of a minimization of a least squares function. The error distributions determined by using all possible points within the LTA volume correspond to the FRE in the surgical volume. When using only the exterior points to register the LTA volume, the remaining interior points position was still known and it is thus possible to compare their actual registered position with the position they should be located in giving a distribution of the TRE within the tested volume; a more realistic scenario for an IGNS intervention since points are generally chosen on the surface of the patient, not throughout the entire surgical volume. The purpose of this portion of the experiment is to determine how different fiducial

configurations affect the distribution of the TRE independent of the FRE and so this is why a large number of fiducial points along the edges are used, making sure that the FRE will reach its plateau value.

Table 3.1: Camera positions

Camera Position Relative to Middle of LTA	Height above LTA
Middle of LTA	50 cm
Middle of LTA	100 cm
45 degrees clockwise	50 cm
45 degrees clockwise	100 cm
80 degrees clockwise	50 cm
80 degrees clockwise	100 cm
45 degrees counterclockwise	50 cm
45 degrees counterclockwise	100 cm
80 degrees counterclockwise	50 cm
80 degrees counterclockwise	100 cm

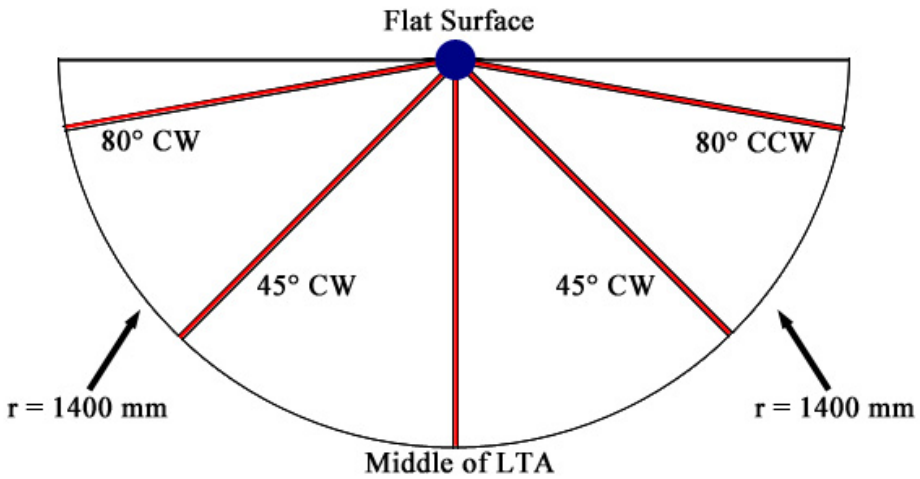

Figure 3.7 Overhead view of camera setup for registration error experiments. The blue circle represents the flat surface that the LTA is on, the black curve represents a distance of 1400 cm from this surface and the intersection of the red line and black curve is where the base of the tripod of the camera was placed. CW represents a clockwise movement and CCW represents a counterclockwise movement, relative to the middle of the LTA.

Chapter 4

Results

In this chapter, a summary of the results obtained from the error quantification and characterization experiments described in the previous chapter is presented. The first section describes the results of the pointer calibration experiments and is followed by the tracking error experiments and finally with the characterization of different registration error metrics.

4.1 Pointer Calibration Errors

A linear regression model was used to fit the calibration error data. This is shown in Figure 4.1. The data for the experiments involving a moving DRF and a static DRF for each camera were plotted on the same axes. A strong positive correlation was found between the calibration RMS and the distance from the camera for a moving DRF, for both pointers. A positive correlation was also found between the calibration RMS and the distance from the camera for a moving DRF. This was evident for both pointers and both cameras except for the NDI lab pointer and camera 1 with a moving DRF, where the calibration RMS stayed relatively constant.

4.2 Tracking Error Measurements

The mean jitter as a function of distance from each camera along the z-axis was plotted for both cameras with both the TA001-133 and TA201-4005 marker tools. The mean of the lowest 90% of the jitter values for each constant z depth was plotted as a function of z as to avoid some of the major

outliers that were in the top 10% of the distribution of jitter values[1].These points were used to find a least-squares fit to a quadratic model to the remaining distribution (Khadem, et al. 2000):

$$J(z) \approx \frac{az^2 + bz + c}{10000} \tag{4.1}$$

Where *z* is measured in meters. These results are summarized in Figure 4.2 which shows the coefficients *a, b* and *c* for each camera-marker pair. The range for which the jitter was calculated corresponded to the calibrated volume of the camera of 1400 mm to 2400 mm. The lowest 90% of the jitter values exhibit an RMS jitter range of 0.30 ± 0.18 for the C1-TA001-133 combination, 0.33 ± 0.21or the C1-Ta201-4005 combination, 0.32 ± 0.22 for the C2-TA001-133 combination and 0.32 ± 0.18 for the C2-Ta201-4005 combination. Table 4.1 summarizes the details of these results. The jitter along the x and y axes were roughly equal and constant on the order of 0.03 mm or less and were insignificant compared to the jitter in the z direction. Figure 4.3 shows the maximum jitter values of the full distribution (without removal of the top 10%).

Table 4.1 Camera-Marker Jitter Value Ranges

Combination	μ (mm)	Δ^+ (mm)	Δ^- (mm)	J_{centre} (mm)	Δ (mm)
C1 + TA001-133	0.25	0.23	-0.13	0.30	0.18
C1 + TA201-4005	0.30	0.24	-0.18	0.33	0.21
C2 + TA001-133	0.24	0.26	-0.14	0.32	0.22
C2 + TA201-4005	0.29	0.22	-0.15	0.32	0.18

Figure 4.3 is a bar graph that shows the maximum jitter values for each camera-marker pair for all the values as well as the maximum jitter values for the lowest 90% of the jitter values.

[1] Reasons for the large amount of outliers are discussed in the following chapter.

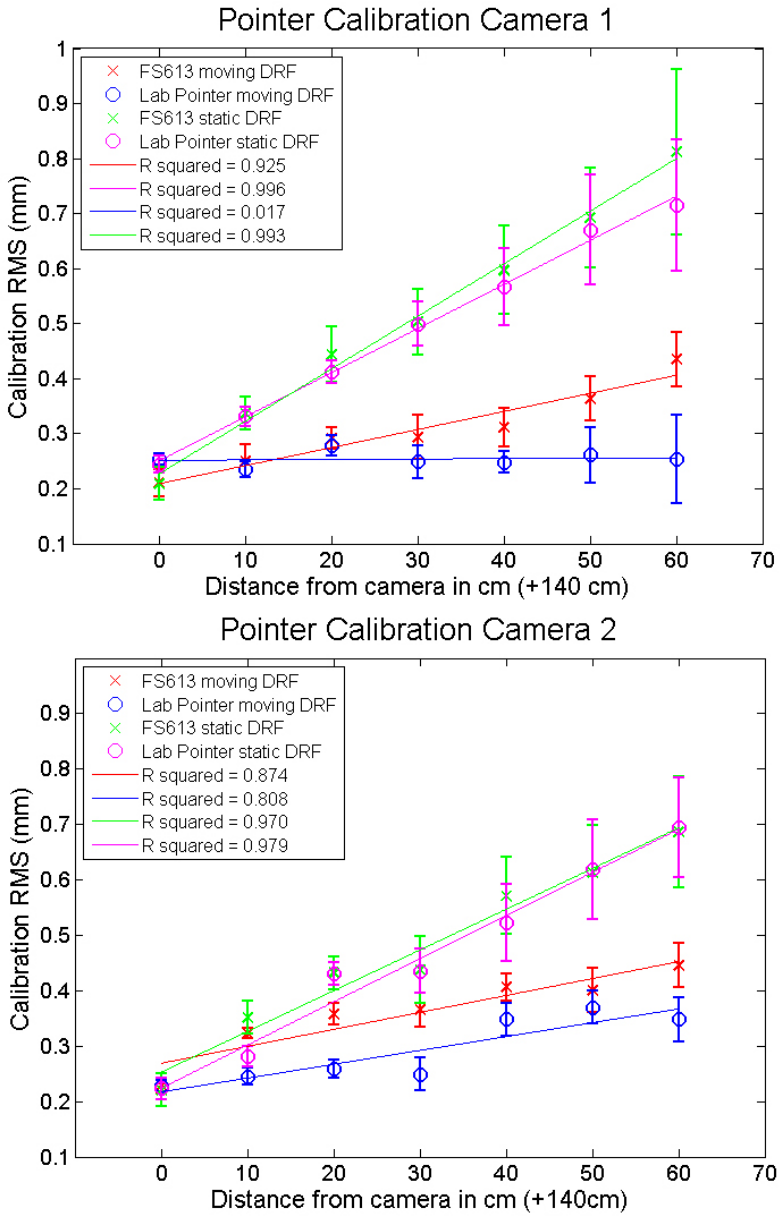

Figure 4.1 Pointer calibration results for both Polaris Cameras. Distance from camera measures the distance from the front of the calibrated viewing volume (See Fig 3.2).

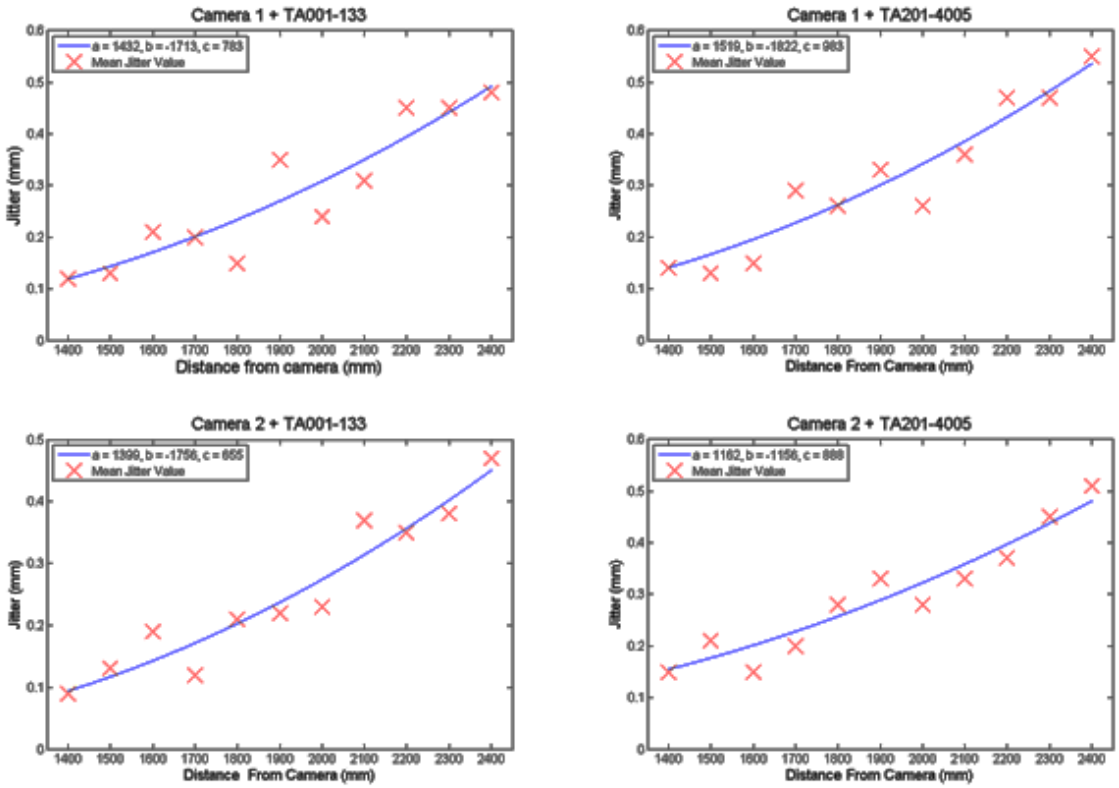

Figure 4.2 Jitter as a function of depth for four different camera-marker combinations.

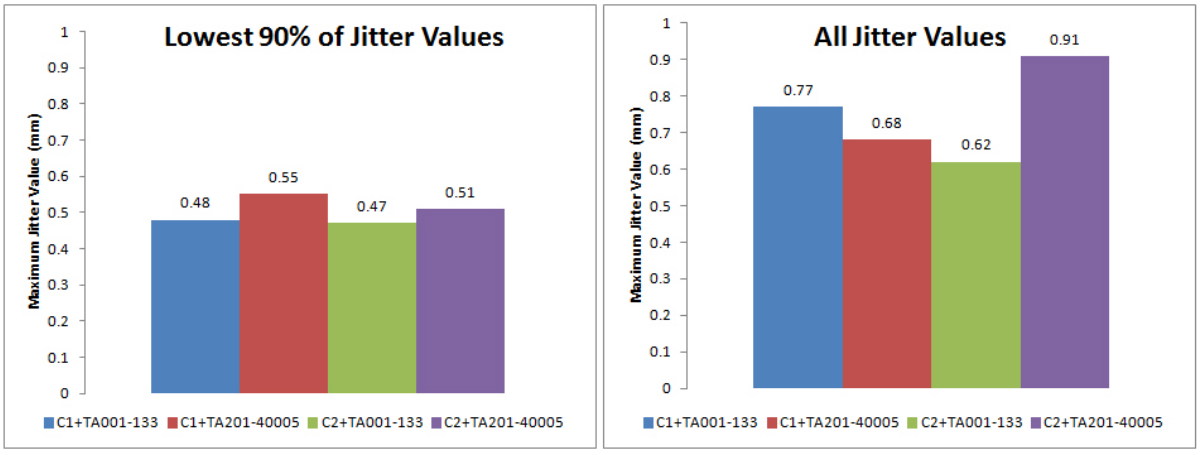

Figure 4.3 Max jitter values for the lowest 90% of the jitter data (left) and all of the jitter data (right).

4.3 Fiducial Registration Error Distributions in the LTA

To evaluate the fiducial registration error in the LTA volume, the 700 points, corresponding to the milled holes (100 holes in the LTA x 7 LTA positions), were used as fiducial markers to register the volume in an image space for 10 different camera positions. This was done for both cameras and both the Traxtal FS613 pointer and NDI lab pointer. The FRE for the FS613-C1 pair was 0.75 mm, the FRE for the FS613-C2 pair was 0.75mm, the FRE for the lab pointer-C1 pair was 0.82 mm and the FRE for the lab pointer-C2 pair was 0.76 mm. Figure 4.4 is a box plot that shows the range, median and max/min of the individual FRE values for each camera-tool pair within the LTA. No statistically significant differences between these distributions were found. Figure 4.5 is also a box plot that shows the variation (standard deviation) of the individual FRE values for each camera-tool pair within the LTA.

The evolution of the mean FRE as a function of number of points used is shown in Figure 4.6 for camera 1 and Figure 4.7 for camera 2. A red ellipse is placed on these graphs to indicate the typical clinical range for number of points used as fiducials for a registration procedure in an IGNS intervention.

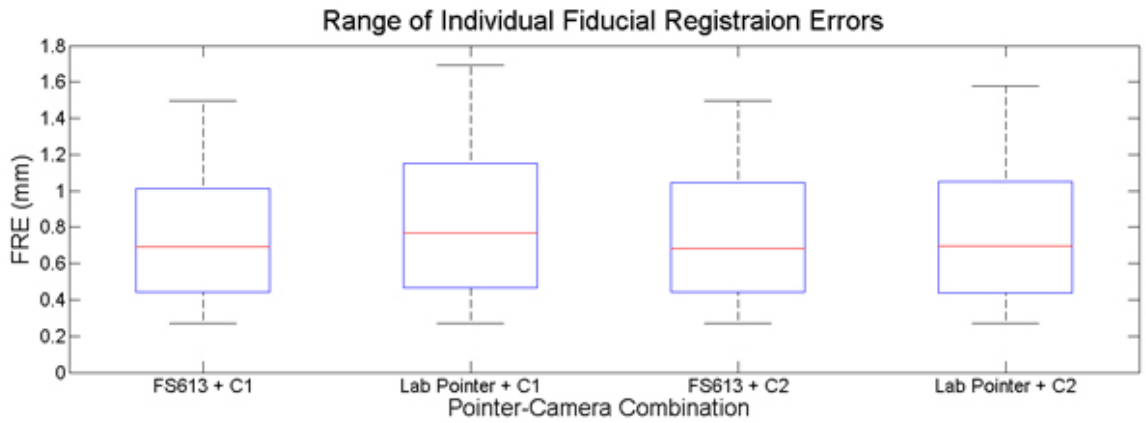

Figure 4.4 Boxplot showing the ranges of the individual FREs. The red line represents the median of the range, the top and bottom edges represent the 25th and 75th percentiles respectively, and the whiskers extend to the most extreme points not considered outliers.

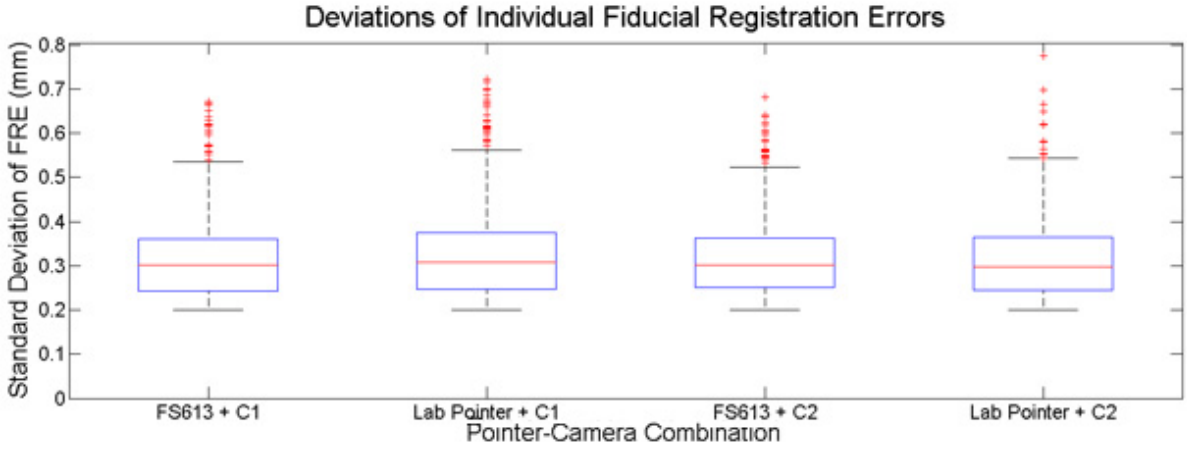

Figure 4.5 Box plot showing the standard deviation of the individual FREs. The red line represents the median, the top and bottom edges of the boxes are the 25th and 75th percentile respectively and the whiskers extend to the maximum and minimum points not considered outliers. The red crosshairs represent points that are considered outliers in the data set.

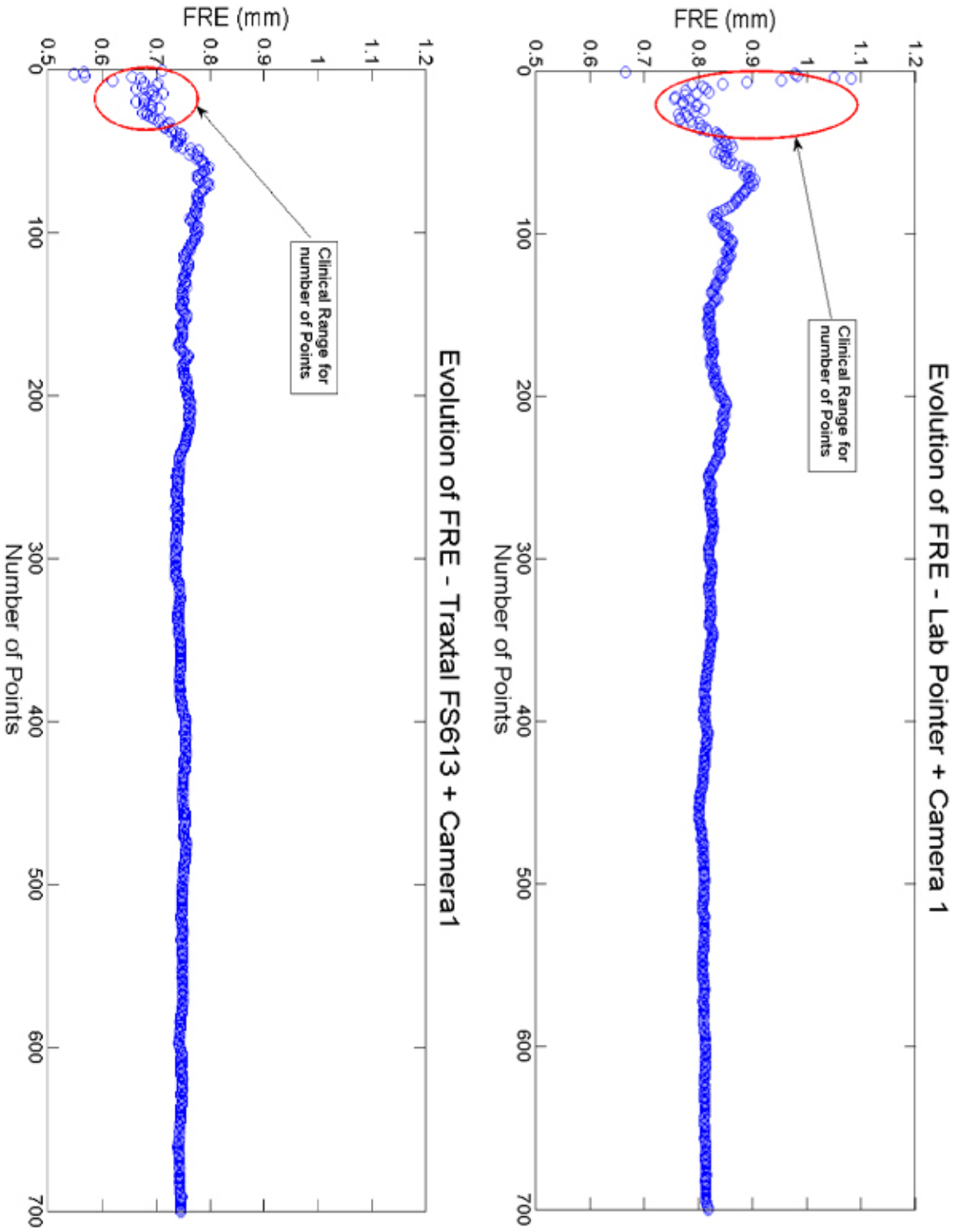

Figure 4.6 Evolution of FRE as a function of number of points used for Camera 1 and both pointers.

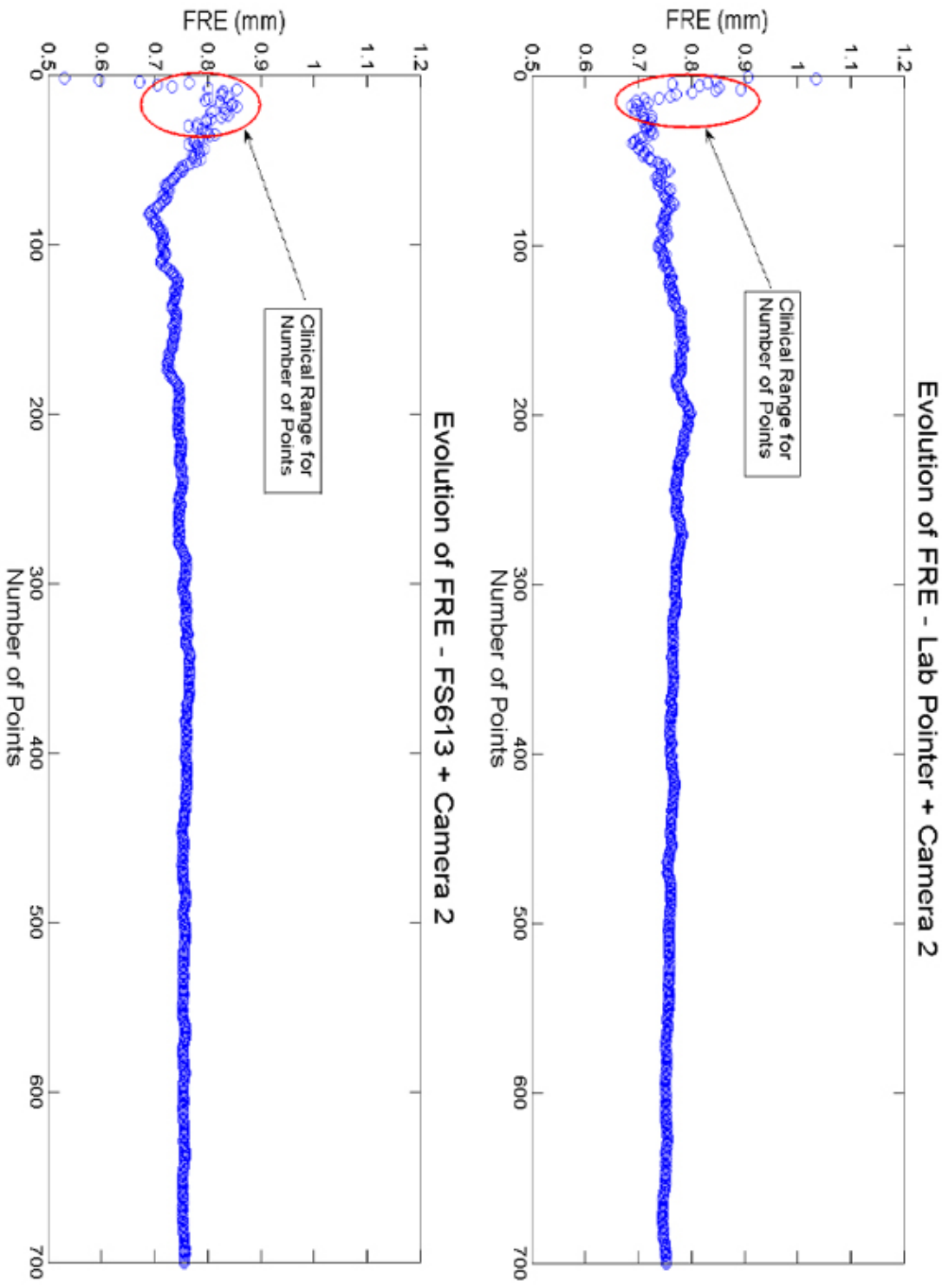

Figure 4.7 Evolution of FRE as a function of number of points used for Camera 2 and both pointers.

4.4 Target Registration Error Distribution in the LTA

From the FRE experiments, it can be seen that the FRE plateaus to a steady value after 100 or more points have been used for the registration for the spatial distribution of points used in these experiments. To obtain a TRE distribution in the LTA, the outer border of each depth, and the outer border plus 2 central points of each depth of the LTA plane were used as fiducial points (corresponding to 252 and 266 points respectfully) for registration of the LTA in image space. Since the position of the remaining 448/434 points is known in physical space and can be determined in the registered image space we can evaluate the TRE at these points. The mean TRE distribution for the central 40mm x 40mm x 500mm volume was evaluated for each camera-pointer tool pair for 10 different camera positions. Figure 4.8 through 4.15 are histograms showing the distribution of the TRE values for exact target locations with the theoretical chi-squared probability distribution overlaid along with an illustrative z-depth slice of the mean and standard deviation of the TRE distribution for each camera-pointer tool pair. Overall, the TRE within the region of interest varies from 0.3 to 1.8 mm.

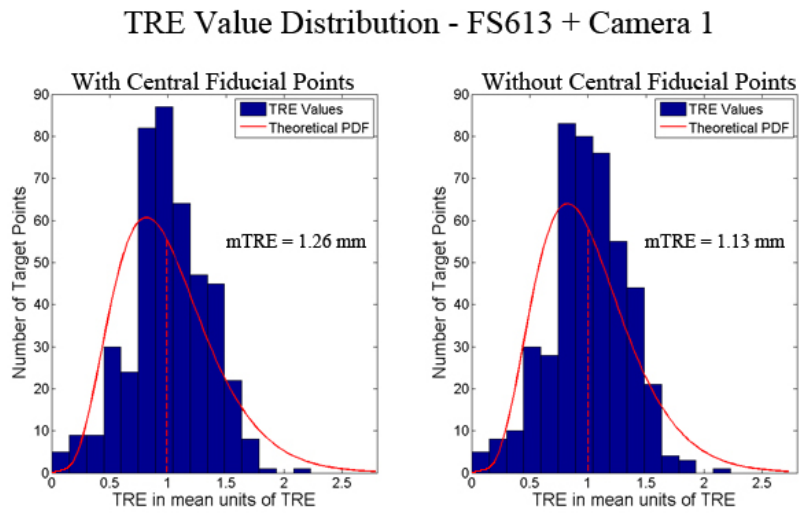

Figure 4.8 Histogram of the TRE value distributions in terms of the mean TRE for the FS613 + Camera 1. The blue bars represent the measured TRE values and the red curve represents the theoretical chi-squared PDF of these values with the red dotted line representing the mean value.

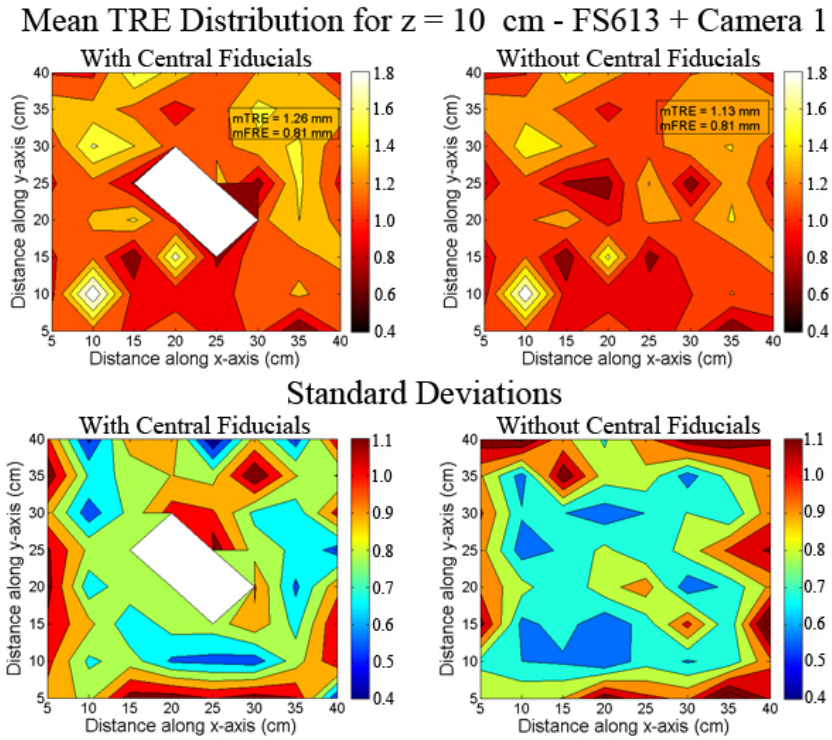

Figure 4.9 Illustrative example of the mean and standard deviation contours for a slice (at a distance of 10cm) of the LTA volume for both sets of fiducial configurations for the Traxtal FS613 pointer and Camera 1. The white rectangle corresponds to where the central fiducials were located.

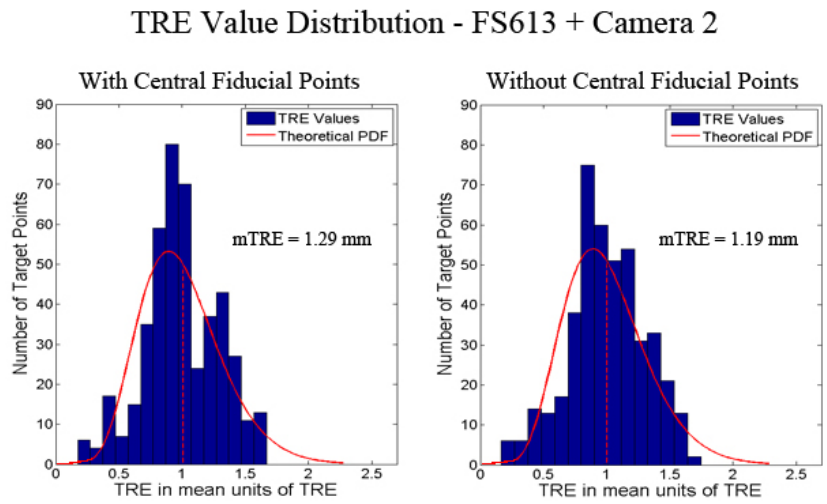

Figure 4.10 Histogram of the TRE value distributions in terms of the mean TRE for the FS613 + Camera 2. The blue bars represent the measured TRE values and the red curve represents the theoretical chi-squared PDF of these values with the red dotted line representing the mean value.

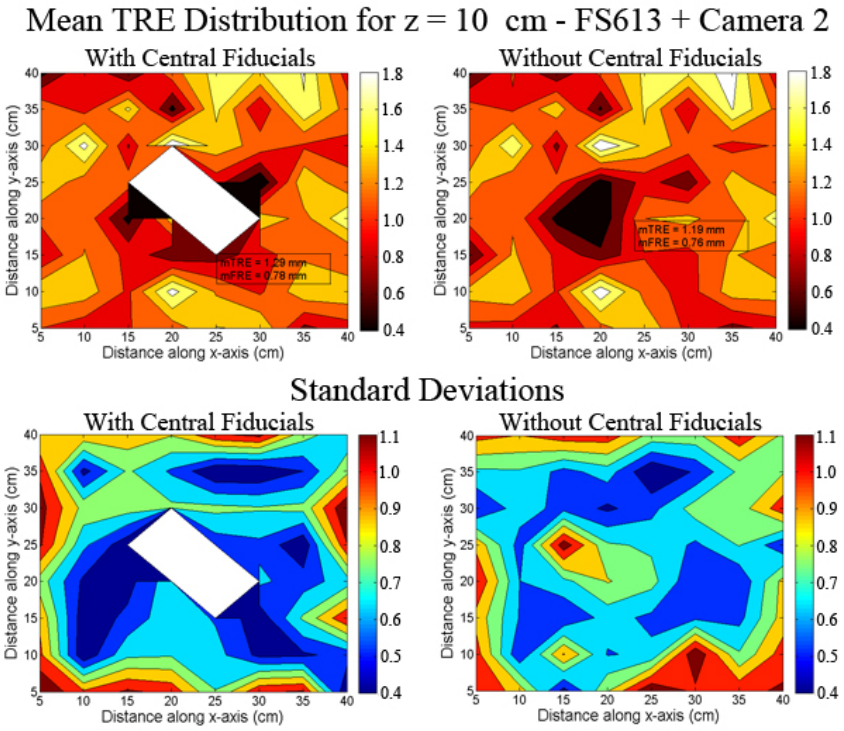

Figure 4.11 Illustrative example of the mean and standard deviation contours for a slice (at a distance of 10cm) of the LTA volume for both sets of fiducial configurations for the Traxtal FS613 pointer and Camera 2. The white rectangle corresponds to where the central fiducials were located.

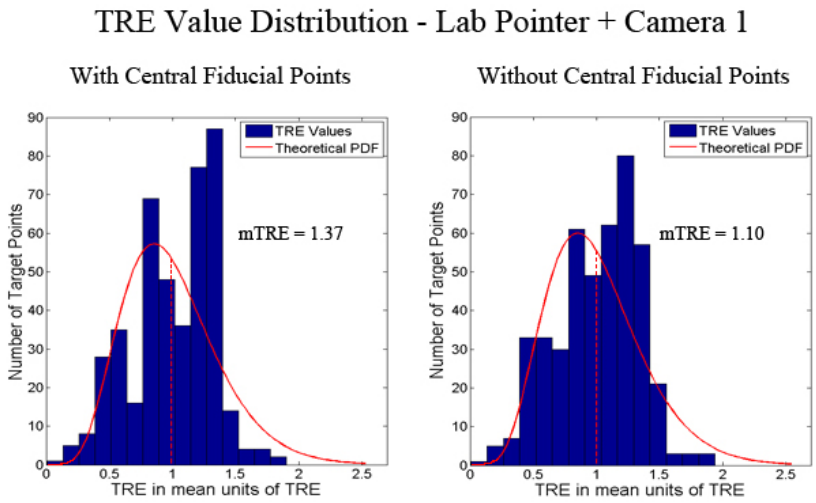

Figure 4.12 Histogram of the TRE value distributions in terms of the mean TRE for the lab pointer + Camera 1. The blue bars represent the measured TRE values and the red curve represents the theoretical chi-squared PDF of these values with the red dotted line representing the mean value.

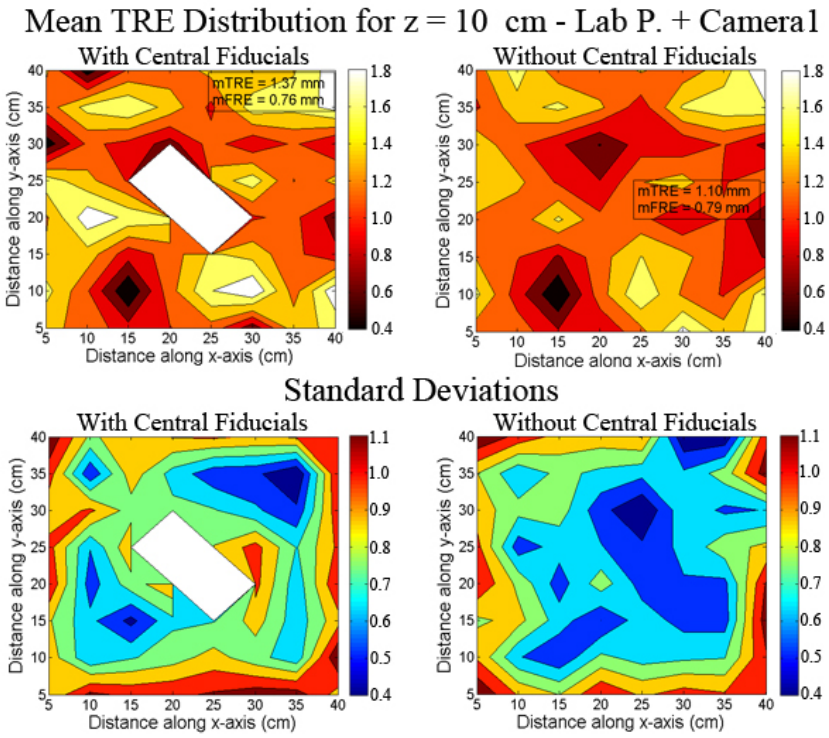

Figure 4.13 Illustrative example of the mean and standard deviation contours for a slice (at a distance of 10cm) of the LTA volume for both sets of fiducial configurations for the NDI lab pointer and Camera 1. The white rectangle corresponds to where the central fiducials were located.

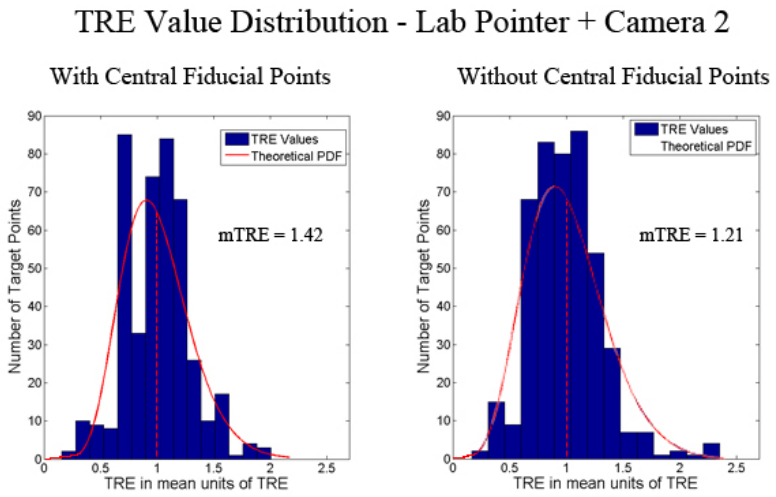

Figure 4.14 Histogram of the TRE value distributions in terms of the mean TRE for the lab pointer and camera 2. The blue bars represent the measured TRE values and the red curve represents the theoretical chi-squared PDF of these values with the red dotted line representing the mean value

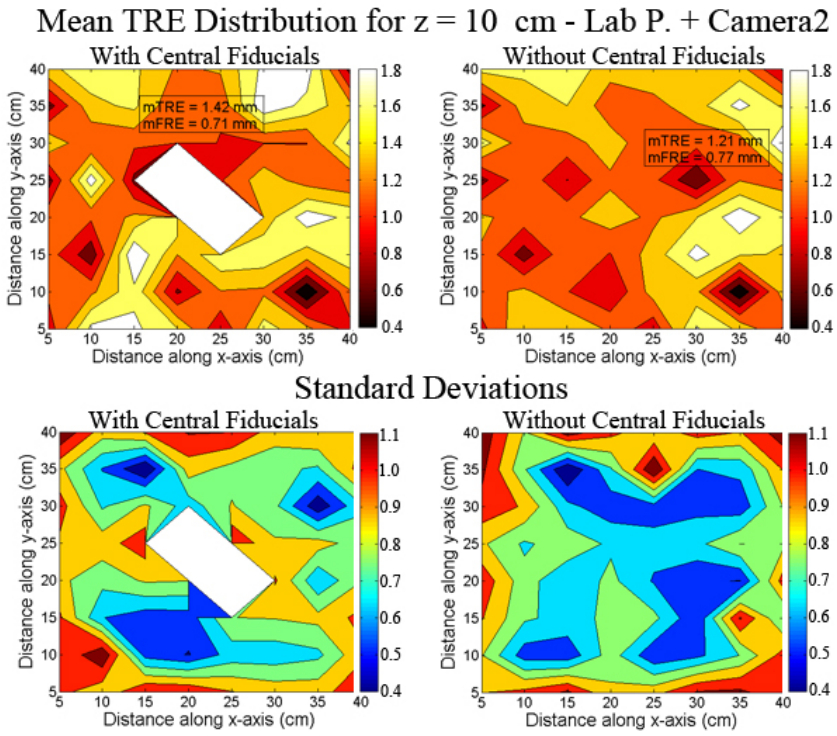

Figure 4.15 Illustrative example of the mean and standard deviation contours for a slice (at a distance of 10cm) of the LTA volume for both sets of fiducial configurations for the NDI lab pointer and Camera 2. The white rectangle corresponds to where the central fiducials were located.

Chapter 5

Discussion

The goal of this thesis was to evaluate and characterize the different errors introduced through the use of tracking technology and tracked tools and their impact on different types of registration errors in the context of IGNS. This was investigated through different experiments that examined different characteristics of each of these components, including the calibration of the tracked pointers, the jitter of the tracking camera and the effects of different types of fiducial arrangements on registration performance and the error associated with it. The following chapter discusses the results obtained in these experiments as well as some of the advantages and disadvantages associated with these results in the context of IGNS.

5.1 Pointer Calibration Errors

Two different surgical pointers were calibrated in a series of experiments to investigate the effect of the distance from the camera and the distance from the DRF on the calibration of the pointer. The calibration of the pointer is an important measure of accuracy for IGNS interventions. It gives an insight into how accurately the physical position of the tracked pointer is being represented in the image space. The factor that had the largest effect on the quality of the calibration was its distance from the DRF; as the pointer's distance from the DRF increased, the calibration RMS also increased. When the pointers were calibrated at deeper distances in the camera's calibrated volume, but at the same distance from the DRF, the calibration remained relatively constant for all distances. As the distance from the camera increased the consistency of the pointer's calibration began to vary

considerably, a finding consistent with the jitter measurements performed previously showing the camera's decreased ability to precisely report tracked objects' position at further distances.

The two pointers used had two different sphere configurations; the Traxtal FS613 had 4 non-planar spheres whereas the NDI pointer had 3 collinear spheres. The difference between camera-pointer calibration pairs was not significantly different. The NDI pointer showed a smaller calibration RMS with smaller standard deviations than the Traxtal pointer for all experiments. This can be explained by geometry and the camera's ability to view each of these shapes. For each tracked tool, there is a special file that explains the configuration of the tool's spheres to the camera. A linear configuration, like that of the NDI pointer, is more easily visible to the camera compared to a nonlinear arrangement like that of the Traxtal pointer, where one of the 4 spheres might hide the others. At any distance and orientation (about the linear configuration) the 3 spheres make a clear and distinct shape that is easy for the camera to see and track. The same can't be said about the nonlinear sphere arrangement on the Traxtal pointer where the position is sometimes more difficult to determine leading to potentially higher RMS if the pointer is not calibrated carefully.

The results from these experiments give another important implication in regards to minimizing error for an IGNS intervention. The tool to be calibrated should be as close to the DRF as possible when being calibrated and also as close to the front of the calibrated camera volume as possible to reduce variation in the calibration. The calibration of the tracked pointer represents another component of the fiducial localization error for registration procedures and minimizing it will help reduce the overall error of the entire IGNS intervention.

5.2 Tracking Error

The jitter for each camera tool pair was dominated by jitter in the z-direction; increasing with increasing distance from the camera. Only the lowest 90% of jitter values were used in the analysis since the upper 10% showed jitter values that were considered outliers and seemed to inappropriately influence the data. Reasons for the large amount of outliers in these experiments is not fully understood, however, similar experiments by Khadem (Khadem, et al. 2000) show that for some optical systems, time varying external disturbances may affect some of the jitter measurements. Every camera-tool pair fit the quadratic model of jitter well with no observable difference between different cameras or different tools used. This trend of increased jitter as a function of distance from the camera leads to an obvious method to improve the precision of tracking during the IGNS

intervention: keep the distance between the tracked tools and the camera at a minimum, and preferably at the front edge of the calibrated volume of the camera. The x and y jitter components were relatively constant and equal in magnitude on the order of 0.03 mm or less and insignificant compared to the jitter values along the z-axis.

This has important implications when designing the camera setup for an IGNS intervention in order to maximize precision and decrease the error associated with tracking. First, the camera should be as close as possible to the surgical field with the orientation of the camera such that the z-direction is most strongly aligned with the surgical direction that is the least clinically significant. Once this has been established, the camera can be rotated about this axis without any significant loss to precision in the surgical volume.

It is important to note that the method used to evaluate the precision of position readings only gives an estimate of the camera's error associated with static measurements. There is generally only one tool in an IGNS intervention that remains static, the DRF, which is important for relating the position of all other tracked tools, however, there are different errors and inaccuracies associated with tracking dynamic targets. Jitter gives a good estimate of the camera's inherent accuracy but the true overall "tracking error" is most likely dependent on more complex parameters. Finally, the jitter measurement considered only positions of a tracked tool that were parallel to the x-y plane. In a typical IGNS intervention, the tools and DRF are not bound to such a constraint and are free to be in any orientation as long as all spheres are visible to the camera which is likely to change the camera's ability to accurately track their position.

The jitter measurement is an important component of one of the least understood type of registration errors, the fiducial localization error, the error associated with physically locating a target in a space. Determining different sources of this type of error enables one to determine ways to minimize it, and thus have a direct impact on improving registration procedures and minimizing error associated with different parts of the IGNS intervention. The camera used for the intervention should also be designed such that it has a calibrated volume that can completely encompass the surgical field in order to avoid changing the camera position.

5.3 Registration Error Distributions in the LTA

The entire LTA grid was used as fiducials to perform a rigid registration between the LTA and a corresponding grid of points in a virtual image for a range of camera positions. The virtual image was

designed as to eliminate all localization errors of fiducials in the image space so that all errors that were measured were due to errors in physical space alone. This is an unrealistic scenario in terms of a true IGNS intervention since many errors can be introduced from locating fiducial points in a corresponding target image as well as on the physical anatomy of a patient, but the goal of this set of experiments was to determine errors solely related to the physical space.

For each camera-pointer tool pair, a similar trend was observed for the mean FRE as the number of points used for the registration increased. The mean FRE starts at an initial value and then slowly converges towards a steady value. This value was approximately 0.75 mm for all sets of camera and pointer tool pairs. This is consistent with the mathematical theory described in the background section, specifically equation 2.12 which expresses, up to the second order, that $\langle FRE^2 \rangle$ is approximately equal to $\langle FLE^2 \rangle$ multiplied by a proportionality factor that tends towards one with an increasing number of fiducials used. The $\langle FRE^2 \rangle$ obtained for these experiments was slightly higher than the combined squared mean values of the two sources of localization errors measured in this thesis meaning that the error that is unaccounted for is due to either error in construction of the LTA or other parameters related to the localization error not investigated in this study. This may include things such as how clean the reflective spheres on the tracked tools are, and thus may justify the manufacturer's suggestion to always use new reflective spheres.

Several important characteristics of the FRE can be seen from the experiments. First, when only a small number of fiducials are used, if any of the initial points have a large (or small) FRE associated with them, the mean FRE tends to be much higher (or lower) than the value that the mean FRE converges to with a high number of fiducials. The variation of the FRE for fiducials located at the edges of the camera volume tended to be higher than those located more in the center of the calibrated volume. As shown by the red ellipse in Figures 4.6 and 4.7, the typical range of fiducials or landmarks used in an IGNS procedure for registration is between 5 – 20. This is the region where the mean FRE varied the most and was the most unstable for all camera-pointer pairs. This means that, when possible, more fiducials should be used for a registration procedure if the FRE is to be minimized. The mean FRE tended to converge after approximately 100 points are used. Using 100 points in an actual IGNS procedure is an unrealistic goal since it takes a lot of time and resources to accurately locate these points on both the physical anatomy of a patient as well as on the preoperative image of a patient, but caution should be taken when choosing this small number of points in order to avoid bias from the mean FRE from outlying large errors. In practice, the FRE for the initial patient-

to-image registration is on the order to 1 mm – 3 mm but can be far worse if the corresponding points are not chosen carefully and if a specific point is far worse than the mean FRE. As outlined above, the impact of these outlying errors can be reduced by using more points. A solution to not using 100 different points may be to use the same point multiple times thus increasing the total number of points used and effectively ensuring that a certain group of corresponding points is properly located.

The second type of error distribution in the LTA observed was the TRE. The two different types of fiducial configurations used for registration showed some slight but interesting differences in the error distributions that arose. Since the goal of this section was to focus on the distribution of the TRE for different fiducial arrangements and equipment and not as a function of the quality of the FRE, a large number of fiducials were used so that the mean FRE converged to the value described in the previous experiment. There was no significant difference between the mean TRE or TRE distribution for different cameras. The mean TRE was lower when using the Traxtal FS613 pointer compared to the NDI lab pointer. The distribution of the TRE values for all the experiments were investigated and the histograms were plotted alongside the theoretical probability distribution described by Labadie, Davis & Fitzpatrick in 2005, and all fit reasonably well with this distribution. However, since these TRE distributions are based on the mean of 10 different data sets the TRE was more highly biased to the mean value of the TRE compared to theoretical prediction.

For all camera-pointer pairs, the configuration of fiducials along only the edge of the LTA, as compared to the configuration comprising the fiducials along the edges with some central points had a smaller mean TRE. For all camera-pointer pairs and fiducial configurations, the standard deviations of the TRE was highest at points closest to fiducials and smallest at points furthest from the fiducial points. The TRE was generally highest at points closest to the fiducials near the edge of the LTA. Some of these results can be explained by lever effects, where the TRE will be best in the centre of the fiducial configuration and increase with distance from the centre. This is seen well in the TRE distribution contours. Also since the value of the FRE around these points changes from camera position to camera position it is more likely that the target points closest to them will suffer similar changes and variations.

5.4 Other Sources of Error

There are many different sources of error that can arise in an IGNS intervention that weren't investigated here that may have a significant impact on the overall registration accuracy. The

experiments here were designed to determine some of the baseline error sources from the technology used in ideal and highly controlled conditions. Biological sources of error were not investigated, as they were beyond the scope of this thesis. In surgery many of these conditions are compromised leading to many other sources of error. Due to the vast amount of equipment needed in the OR, the tracking camera cannot always be placed so that the patient is at the front of the calibrated volume. In turn, as we've shown here this has an approximately quadratic effect on the camera's ability to properly report the position of the tracked tools. When acquiring preoperative images of the patients, this can occur days to weeks before the actual surgery which could reflect an anatomical difference between the two volumes and thus invalidate the assumption that a rigid body transformation is sufficient to describe the image-to-patient mapping. Landmark points for registration in the OR are based on anatomical points on the skin of the patient. The skin is a very flexible organ and may not be in the same position when the preoperative images were taken compared to when the patient is in the OR. This type of error may also be amplified by MR susceptibility artifacts during imaging, or if the patient is overweight. A recent study investigating the loss of accuracy in the OR for IGNS interventions has grouped errors into three different categories, technical, operational and biological. The most important physical aspect is the errors arising from patient-image registration (Stieglitz, et al. 2013), which was also investigated in this work. The work by Stieglitz also showed that the use of different surgical equipment, such as surgical drapes, had a major effect on the accuracy of the IGNS intervention. The weight of the drapes affect the clamp, the patient's head and any instruments attached to it as well. These drapes are generally stretched and straightened to obtain a big sterile field, which enhances the forces exerted on the patient and the increasing weight of the drapes when soaked through irrigation liquid and blood increase this effect (Stieglitz, et al. 2013).

As the length of the surgery increases, there are many biological factors that can affect the accuracy of the IGNS intervention. The main one is the phenomenon of brain shift that occurs throughout surgery. Changes in blood flow, CSF volume, brain volume and the use of different steroids and other drugs during surgery will cause the brain volume to change and deviate from the preoperative images of the anatomy thus compromising the accuracy of the patient-image registration and neuronavigation almost entirely.

Chapter 6 Conclusion and Future Work

IGNS interventions are complex procedures involving many different technologies and tools that can lead to many different sources of error. Before surgery even begins, there are many factors that can affect neuronavigation accuracy; the position of the infrared tracking camera can have a strong impact on the camera's ability to properly and consistently report the position of any tracked tools. This tracking error can be on the order of 0.5 mm if care is not taken to properly position the patient as close to the calibrated volume of the camera as well with the most clinically significant direction not normal to the plane of the camera. The calibration of the different tools used for the IGNS procedure can have a large impact on the system's ability to accurately report a physical location on a patient to the corresponding location on an image. The tools being used should be calibrated at the front of the calibrated volume of the camera and as close to the reference tracking tool as possible. Even when this is done, the calibration error can be on the order of 0.3 mm to 0.8 mm. These two types of error are generally associated in the category of fiducial localization errors. Choosing corresponding points on a physical object or patient as well as on the image of interest is very important in obtaining a good registration between the two spaces. Using as many points as possible will decrease the overall FRE until it reaches a threshold value determined by the errors associated in localizing these points. Since it is generally impossible to choose a large number of anatomical landmarks in a real IGNS intervention where the mean TRE plateaus, a future study will focus on how choosing the same

landmark point pairs multiple times will affect the FRE and its distribution. The error distribution among target points not used as fiducials generally follows a chi-squared distribution and shows a higher value of error with a larger variation for points that farther from the center of the region of interest. Fiducials points should thus be chosen to surround the surgical volume of interest, while also leaving sufficient distance so that the registration error within these points is smallest where the surgical procedure is to occur. Many characteristics of different types of registration errors, including the FLE, FRE, and TRE, were investigated in this work and were observed to agree well with some of the theoretical descriptions of their behavior. There are many other sources of error that are unavoidable in IGNS interventions including different surgical equipment used as well as biological phenomena that occur due to brain movement so care and caution must be taken to minimize error in all aspects where it is possible.

In the future, the difference between the trends of the TRE and FRE under idealistic conditions using a precision milled LTA will be compared to clinical cases of IGNS lesions to investigate some of the main differences that occur between the characteristics of these types of error from the lab to the clinic and to try and determine other sources that may contribute to this difference in error. Also, a study to quantify the amount of brain shift that occurs throughout surgery and to try and correct these shifts using intraoperative ultrasound will be performed as part of a doctorate degree.

Bibliography

Brinker, T., G. Arango, and J. Kaminsky. “An experimental approach to imaged guided skull base surgery employing a microscope-based neuronavigation system.” *acta neurochirurgica (Wien)* 140, no. 9 (1998): 883-889.

Comeau, R. M., A. F. Sadikot, A. Fenster, and T. M. Peters. “Intraoperative ultasound for guidance and tissue shift correction in image-guided neurosurgery.” *Medical Physics* 27, no. 4 (2000): 787 - 800.

Cuchet, E., J. Knoplioch, D. Dormont, and C. Marsault. “Registration in Neurosurgery and Neuroradiotherapy Applications.” *Journal of Imaged Guided Surgery* 1 (1995): 198-207.

Day, J. S., D. J. Murdoch, and G. A. Dumas. “Calibration of position and angular data from a magnetic tracking device.” *Journal of Biomechanics* 33, no. 8 (2000): 1039-1045.

Fitzpatrick, J. M. “Fiducial registration error and target registration error are uncorrelated.” *Proceedings of SPIE.* 2009. 1 - 12.

Fitzpatrick, J. M., and J. B. West. “The Distribution of Target Registration Error in Rigid-Body Point-Based Registration.” *IEEE T Med Imaging* 20, no. 9 (2001): 917 - 927.

Fitzpatrick, J. M., J. B. West, and C. R. Maurer Jr. “Predicting Error in Rigid-Body Point-Based Registration.” *IEEE T Med Imaging* 17, no. 5 (1998): 694 - 702.

Fonov, V., et al. “Improved precision in the measurement of longitudinal global and regional volumetric changes via a novel MRI gradient distortion characterization and correction technique.” *Lecture Notes in Computer Science* 6326 (2010): 324 - 333.

Germano, I. M., H. Villalobos, A. Silvers, and K. D. Post. “Clinical use of the optical digitizer for intracranial neuronavigation.” *Neurosurgery* 45, no. 2 (1999): 261-269.

Golfinos, J. G., B. C. Fitzpatrick, L. R. Smith, and R. F. Spetzler. “Clinical use of a frameless stereotactic arm: results of 325 cases.” *Journal of Neurosurgery* 83, no. 2 (1995): 197-205.

Gumprecht, H. K., D. C. Widenka, and C. B. Lumenta. “BrainLab VectorVision Neuronavigation System: technology and clinical experiences in 131 cases.” *Neurosurgery* 25, no. 1-2 (2002): 797-803.

Hartov, A., S. D. Eisner, D. W. Roberts, K. D. Paulsen, L. A. Platenik, and M. I. Miga. “Error analysis for a free-hand three-dimensional ultrasound system for neuronavigation.” *Neurosurical Focus* 6, no. 3 (1999): Article 5.

Hartov, A., S. D. Eisner, D. W. Roberts, K. D. Paulsen, L. A. Platenik, and M. I. Miga. “Error analysis for a free-hand three-dimensional ultrasound system for neuronavigation.” *Neurosurg Focus* 6, no. 3 (1999).

Hatiboglu, Mustafa A., et al. “Impact o Intraoperative High-Field Magnetic Resonance Imaging Guidance on Glioma Surgery: A Prospective Volumetric Analysis.” *Neurosurgery* 64, no. 6 (2009): 1073-1081.

Hill, D. L. G., P. G. Batchelor, M. Holden, and D. J. Hawkes. “Medical Image Registration.” *Physics in Medicine and Biology* 46, no. 3 (2001).

Hill, Derek L. G., Calvin R. Maurer, Robert J. Maciunas, John A. Barwise, Michael J. Fitzpatrick, and Matthew Y. Wang. “Measurement of Intraoperative Brain Surface Deformation under a Craniotomy.” *Neurosurgery* 43, no. 3 (1998): 514-526.

Khadem, Rasool, et al. “Comparative Tracking Error Analysis of Five Different Optical Tracking Systems.” *Computer Aided Surgery* 5 (2000): 98-107.

Labadie, R. F., B. M. Davis, and J. M. Fitzpatrick. “Image-guided surgery: what is the accuracy.” *Current opinion in otolaryngology & Head and Neck Surgery* 13, no. 1 (2005): 27 - 31.

Lauterbur, P. C. “Image formation by induced local interactions: Examples employing nuclear magnetic resonance.” *Nature* 242 (1973): 190-191.

Liu, W., H. Ding, H. Han, Q. Xue, Z. Sun, and G. Wang. “The Study of Fiducial Localization Error of Image in Point-based Registration.” *31st Annual International Conference of the IEEE EMBS.* 2009. 5088 - 5091.

Maesawa, Satoshi, et al. “Clinical Indications for High-Field 1.5 T Intraoperative Magnetic Resonance Imaging and Neuro-navigation for Neurosurgical Procedures.” *Neurol Med Chir(Tokyo)* 49 (2009): 340-350.

Maintz, J. B., and M. A. Viergever. “A survey of medical image registration.” *Medical Imaging Analysis* 2, no. 1 (1998): 1-36.

Maurer Jr., C. R., et al. “Effect of geometrical distortion correction in MR on image registration accuracy.” *J. Comp. Assisted Tomography* 20 (1996): 666 - 679.

Maurer Jr., Calvin R., et al. “Effect of geometrical distortion correction in MR on image registration accuracy.” *Journal of Computer Assisted Tomography* 20 (1996): 666-679.

McGirt, Matthew J., et al. “Independent association of extent of resection with survival in patients with malignant brain astrocytoma.” *J Neurosurg* 110, no. 1 (2009): 156 - 162.

Mercier, L. “Ultrasound guided brain tumour resection.” Thesis, Computer Engineering, Mcgill University, Montreal, 2011.

Mercier, L., et al. “New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: System description and validation.” *Int J CARS* 6, no. 4 (2011): 507 - 522.

Mistry, D., and A. Banerjee. “Review: Image Registration.” *IJGIP* 2, no. 1 (2012): 18 - 22.

Moghari, M. H., and P. Abolmaesumi. “Understanding the Effect of Bias in Fiducial Localization Error on Point-Based Rigid-Body Registration.” *IEEE T Med Imaging* 29, no. 10 (2010): 1730 - 1738.

Nimsky, Christopher, et al. “Updating Navigation With Intraoperative Image Data.” *Topics in Magnetic Resonance Imaging* 19, no. 4 (2008): 197-204.

O'Donnell., M., A. R. Skovoroda, B. M. Shapo, and S. Y. Emelianov. “Internal Displacement and Strain Imaging Using Ultrasonic Speckle Tracking.” *IEEE Transactions on Ultrasonics, Ferroelectrics, and frequency control* 41, no. 3 (1994): 314-325.

Pfisterer, W. K., S. Papadopoulos, D. A. Drumm, K. Smith, and M. C. Preul. “Fiducial versus nonfiducial neuronavigation registration assessment and considerations of accuracy.” *Neurosurgery* 62, no. 3 (2008): 201-207.

Pillai, P., S. Sammet, and M. Ammirati. “Application accuracy of computed tomography-based, image-guided navigation of temporal bone.” *Neurosurgery* 63, no. 4 suppl 2 (2008): 326-333.

Roberts, David W., Alexander Hartov, Francis E. Kennedy, Michael I. Miga, and Keith D. Paulsen. “Intraoperative Brain Shift and Deformations: A Quantitative Analysis of Cortical Displacement in 28 Cases.” *Neurosurgery* 43, no. 4 (1998): 749-758.

Sagi, H. C., R. Manos, R. Benz, N. R. Ordway, and P. J. Connolly. “Electromagnetic Field-Based Image-Guided Spine Surgery Part One: Results of a Cadaveric Study Evaluating Lumbar Pedicle Screw Placement.” *Spine* 28, no. 17 (2003): 2013 - 2018.

Sibson, R. “Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling.” *J.R. Stat. Soc.: Series B* 41, no. 2 (1979): 217 - 229.

Sipos, E. P., S. A. Tebo, S. J. Zinreich, D. M. Long, and H. Brem. “In vivo accuracy testing and clinical experience with the ISG Viewing Wand.” *Neurosurgery* 39, no. 1 (1996): 194-202.

Stieglitz, L. H., et al. “The silent loss of neuronavigation accuracy: A systematic retrospective analysis of factors influencing the mismatch of frameless stereotactic systems in cranial neurosurgery.” *Neurosurgery* 72 (2013): 796-807.

Thompson, E. M., G. J. Anderson, C. M. Roberts, M. A. Hunt, and N. R. Selden. “Skull-fixated fiducial markers improve accuracy in staged frameless stereotactic epilepsy surgery in children.” *Journal of Neurosurgery: Pediatrics* 7, no. 1 (2011): 116-119.

van den Elsen, P. A., E. J. D. Pol, and M. A. Viergever. “Medical image matching- a review with classification.” *Engeineering in Medicine and Biology Magazine, IEEE* 12, no. 1 (1993): 26-39.

West, J. B., J. M. Fitzpatrick, J. Michael, S. A. Toms, C. R. Maurer, and R. J. Maciunas. “Fiducial Point Placement and the Accuracy of Point-based, Rigid Body Registration.” *Neurosurgery* 48, no. 4 (2001): 810 - 817.

Wiles, A. D., D. G. Thompson, and D. D. Frantz. “Accuracy assessment and interpretation for optical tracking systems.” *SPIE Proceedings.* 2004.

Zitova, B., and J. Flusser. “Image Registration methods: a survey.” *Image and Vision computing* 21, no. 11 (2003): 977-1000.