Mobile and Low-cost Hardware Integration in Neurosurgical Image-Guidance

Étienne Léger

A Thesis
in
The Department
of
Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Computer Science) at
Concordia University
Montréal, Québec, Canada

July 2021

© Étienne Léger, 2021

Concordia University
School of Graduate Studies

This is to certify that the thesis prepared

By: **Étienne Léger**

Entitled: **Mobile and Low-cost Hardware Integration in Neurosurgical Image-Guidance**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Computer Science)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. Chun Wang

_____ External Examiner
Dr. Catherine Laporte

_____ External to Program
Dr. Hassan Rivaz

_____ Examiner
Dr. Charalambos Poullis

_____ Examiner
Dr. Tristan Glatard

_____ Thesis Supervisor
Dr. Marta Kersten-Oertel

_____ Co-supervisor
Dr. Tiberiu Popa

Approved by _____
Dr. Hovhannes Harutyunyan, Graduate Program Director

September 7th, 2021 _____
Dr. Mourad Debbabi, Dean
Gina Cody School of Engineering and Computer Science

Abstract

Mobile and Low-cost Hardware Integration in Neurosurgical Image-Guidance

Étienne Léger, Ph.D.

Concordia University, 2021

It is estimated that 13.8 million patients per year require neurosurgical interventions worldwide, be it for a cerebrovascular disease, stroke, tumour resection, or epilepsy treatment, among others. These procedures involve navigating through and around complex anatomy in an organ where damage to eloquent healthy tissue must be minimized. Neurosurgery thus has very specific constraints compared to most other domains of surgical care. These constraints have made neurosurgery particularly suitable for integrating new technologies. Any new method that has the potential to improve surgical outcomes is worth pursuing, as it has the potential to not only save and prolong lives of patients, but also increase the quality of life post-treatment. In this thesis, novel neurosurgical image-guidance methods are developed, making use of currently available, low-cost off-the-shelf components. In particular, a mobile device (*e.g.,* smartphone or tablet) is integrated into a neuronavigation framework to explore new augmented reality visualization paradigms and novel intuitive interaction methods. The developed tools aim at improving image-guidance using augmented reality to improve intuitiveness and ease of use. Further, we use gestures on the mobile device to increase interactivity with the neuronavigation system in order to provide solutions to the problem of accuracy loss or brain shift that occurs during surgery. Lastly, we explore the effectiveness and accuracy of low-cost hardware components

(*i.e.,* tracking systems and ultrasound) that could be used to replace the current high cost hardware that are integrated into commercial image-guided neurosurgery systems. The results of our work show the feasibility of using mobile devices to improve neurosurgical processes. Augmented reality enables surgeons to focus on the surgical field while getting intuitive guidance information. Mobile devices also allow for easy interaction with the neuronavigation system thus enabling surgeons to directly interact with systems in the operating room to improve accuracy and streamline procedures. Lastly, our results show that low-cost components can be integrated into a neurosurgical guidance system at a fraction of the cost, while having a negligible impact on accuracy. The developed methods have the potential to improve surgical workflows, as well as democratize access to higher quality care worldwide.

Acknowledgments

Thanks to Marta, my mentor.

Merci à Gabrielle, mon pilier.

Thanks to Tiberiu, Simon, Houssem, Jonatan, Louis, Ian, Jeffery, David, Reid, Terry, Xiao and everyone else who helped shape and produce this thesis.

Contributions of Authors

I am the first author of all five (including addendum to Chapter 5) manuscripts included in this dissertation and as such have performed all of the methodological developments, experimental design, data collection and analysis of results. The contributions of co-authors include supervision, software development, technical discussions and review of manuscripts. The following list summarizes the contributions of each author, organized by manuscript:

Chapter 3: Quantifying attention shifts in augmented reality image-guided neurosurgery. *Healthcare Technology Letters*, 4(5), 188–192, 2017.

- Authors: Étienne Léger, Simon Drouin, D. Louis Collins, Tiberiu Popa, Marta Kersten-Oertel
- Contributions: Guarantors of integrity of the study: all authors; study and design concepts: É.L., D.L.C. and M.K.-O.; software development: É.L. and

S.D.; data collection: É.L.; data preparation and analysis: É.L.; supervision: T.P. and M.K.-O.; manuscript preparation: É.L.; manuscript revision: all authors; editing and final version: É.L. and M.K.-O.

Chapter 4: Gesture-based registration correction using a mobile augmented reality image-guided neurosurgery system. *Healthcare Technology Letters*, 5(5), 137–142, 2018.

- Authors: Étienne Léger, Jonatan Reyes, Simon Drouin, D. Louis Collins, Tiberiu Popa, Marta Kersten-Oertel

- Contributions: Guarantors of integrity of the study: all authors; study and design concepts: É.L., S.D., D.L.C. and M.K.-O.; software development: É.L. and J.R.; data collection: É.L.; data preparation and analysis: É.L.; supervision: T.P. and M.K.-O.; manuscript preparation: É.L.; manuscript revision: all authors; editing and final version: É.L. and M.K.-O.

Chapter 5: MARIN: an open-source mobile augmented reality interactive neuronavigation system. *International Journal of Computer Assisted Radiology and Surgery*, 15(6), 1013–1021, 2020.

- Authors: Étienne Léger, Jonatan Reyes, Simon Drouin, Tiberiu Popa, Jeffery A. Hall, D. Louis Collins, Marta Kersten-Oertel

- Contributions: Guarantors of integrity of the study: all authors; study and design concepts: É.L. and M.K.-O.; software development: É.L. and J.R.; data collection in operating room: É.L., S.D. and J.A.H.; data collection in lab and phantom experiment: É.L.; data preparation and analysis: É.L.; supervision: T.P. and M.K.-O.; manuscript preparation: É.L.; manuscript revision: all authors; editing and final version: É.L. and M.K.-O.

Addendum to Chapter 5: A prototype for hemodynamic analysis in neurosurgery implemented on a mobile augmented reality system. *International Journal of Computer Assisted Radiology and Surgery*, 15 (Supplement 1), 91-92, 2020.

- Authors: Étienne Léger, Reid F. Vassallo, Tiberiu Popa, Terry Peters, Yiming Xiao, Marta Kersten-Oertel

- Contributions: Guarantors of integrity of the study: all authors; study and design concepts: É.L. and M.K.-O.; software development: É.L. and R.F.V.; lab experiment: É.L.; data preparation and analysis: É.L.; supervision: T.P. and M.K.-O.; manuscript preparation: É.L.; manuscript revision: all authors; editing and final version: all authors.

Chapter 6: Evaluation of low-cost hardware alternatives for 3D freehand ultrasound reconstruction in image-guided neurosurgery (accepted to the Workshop of Advances in Simplifying Medical UltraSound (ASMUS))

- Authors: Étienne Léger, Houssem Eddine Gueziri, Tiberiu Popa, D. Louis Collins, Marta Kersten-Oertel

- Contributions: Guarantors of integrity of the study: all authors; study and design concepts: É.L., H.E.G. and M.K.-O.; software development: É.L.; data collection: É.L., H.E.G; data preparation and analysis: É.L., H.E.G; supervision: T.P. and M.K.-O.; manuscript preparation: É.L.; manuscript revision and editing of final version: É.L., H.E.G and M.K.-O.

Contents

List of Figures **xiv**

List of Tables **xvii**

List of Abbreviations **xviii**

1 Introduction **1**

1.1 Motivation 4

1.2 Objectives and Contributions 5

1.3 Outline 7

2 Background **8**

2.1 Image-guided Neurosurgery 8

2.1.1 History 9

2.1.2 IGNS System Components 13

2.2 Accuracy of IGNS Systems and Brain Shift 18
2.2.1 Causes of Brain Shift 20
2.2.2 Brain Shift Compensation 23
2.2.3 Intraoperative Imaging for Brain Shift Correction 23
2.3 Visualization and Interaction with Guidance Information 25
2.3.1 Augmented Reality Image-guided Neurosurgery 27

3 Quantifying attention shifts in augmented reality image-guided neurosurgery 36

3.1 Introduction 38
3.2 Related Work 40
3.2.1 AR in Image-guided Neurosurgery 41
3.2.2 Attention Shifts in Surgery 42
3.3 Methodology 43
3.4 Experiment 44
3.5 Results 47
3.5.1 Questionnaire Results 49
3.6 Discussion 51
3.7 Conclusions 54

4 Gesture-based registration correction using a mobile augmented reality image-guided neurosurgery system 55

4.1 Introduction 57
4.2 Related Work 60
4.3 System Description 62
4.4 Methodology 65
4.5 Experiment 66

4.6 Results 69
4.7 Discussion 72
4.8 Conclusion 76

5 MARIN: an open-source mobile augmented reality interactive neuronavigation system 78

5.1 Introduction 80
5.2 Previous Work 82
5.3 System Description 83
5.3.1 Hardware 84
5.3.2 Software 85
5.3.3 MARIN Interface 87
5.4 User Study 88
5.4.1 Phantom Construction 88
5.4.2 Task Description 90
5.4.3 Experimental Procedure 91
5.5 User Study Results 92
5.5.1 Quantitative Measurements 92
5.5.2 Qualitative Feedback 94
5.6 Expert Feedback 95
5.7 Discussion 96
5.8 Conclusion 98

A prototype for video-based hemodynamic analysis in neurosurgery implemented on a mobile augmented reality system 101

5.9 Purpose 101
5.10 Methods 102

5.11 Results 104
5.12 Conclusions 104

6 Evaluation of low-cost hardware alternatives for 3D freehand ultrasound reconstruction in image-guided surgery 106

6.1 Introduction 108
6.2 Previous Work 111
6.3 Methodology 113
6.3.1 Hardware Setups 113
6.3.2 Phantoms and Marker Construction 113
6.3.3 Experiments 115
6.4 Results 119
6.4.1 US Reconstruction Accuracy Results 120
6.4.2 Qualitative Results 122
6.4.3 Preoperative Volume Registration Results 123
6.5 Discussion 124
6.6 Conclusion and Future Work 127

7 Conclusion and future work 129

7.1 Summary of Findings 130
7.1.1 Augmented Reality in Neurosurgery: Cognitive Load 130
7.1.2 Gesture-based Registration Correction 132
7.1.3 Augmented Reality in Neurosurgery: Accuracy 133
7.1.4 Intraoperative Augmented Reality Video-Based Hemodynamic Analysis 134
7.1.5 Low-cost Tracking for 3D Freehand Ultrasound Reconstruction 135
7.2 Future Research Directions 136

7.2.1 Visualization 136
7.2.2 Uncertainty 137
7.2.3 System Validation 137
7.2.4 Mobile AR 138
7.3 Outlook 139

Bibliography **140**

Appendix A User Study Questionnaires **153**
A.1 Focus Shift Study Questionnaire 153
A.2 Manual Registration Correction Assessment Study Questionnaire . . . 165
A.3 Marin Usability Assessment Study Questionnaire 171

List of Figures

1 A typical surgical scene where an IGNS system is used 3

2 Horsley and Clarke's stereotactic frame 10

3 Roberts *et al*'s microscope augmented reality view 12

4 Typical surgical environment when using a traditional neuronavigation system 14

5 Schematic of the registration procedure 17

6 Screenshot of a typical IGNS navigational interface 19

7 Successive MRI acquisitions taken during surgery showing the effect of brain shift 20

8 Example of camera-based AR 27

9 Our developed mobile AR system displaying CTA acquired vessels over a phantom head 32

10 Table summarizing existing mobile AR neuronavigation applications . 34

11 Example of craniotomy planning 40

12 Screenshot of the IGNS monitor view for the Traditional Nav condition for the focus shift experiment 45
13 Experimental set-up for the focus shift experiment 46
14 Boxplots of the total times taken per condition for the focus shift experiment 49
15 Boxplots of the number of attention shifts per condition for the focus shift experiment 50
16 Boxplots of the ratio of time looking at desktop/mobile over total time per condition for the focus shift experiment 51
17 NASA TLX individual scales per condition for the focus shift experiment 52
18 System set-up for gesture control experiment: phantom and AR view 59
19 System set-up for gesture control experiment: tracker, workstation and mobile device 63
20 Screenshot of the iPad AR application during use 65
21 Example registration correction 67
22 Final RMS error against initial misregistration for gesture control experiment 70
23 Final RMS error presented as a percentage of initial misregistration for gesture control experiment 71
24 System set-up for the MARIN evaluation experiment 83
25 Block diagram representing interactions between system components and data flow in MARIN 86
26 Screenshots of MARIN in two different modes of operation 87
27 Experimental phantom design for the MARIN evaluation experiment 89
28 Boxplots showing final distance to target and total time taken, compared per system for the MARIN evaluation experiment 94

29 NASA TLX results for the MARIN evaluation experiment 95
30 Initial tests of MARIN in a clinical setting 96
31 Test setup for hemodynamic experience 103
32 Depiction of trackers and probes used for acquisition in the low-cost reconstruction experiment 114
33 Picture of the experimental wire phantom for low-cost hardware evaluation 115
34 Picture of tracker arrangement for data acquisition for low-cost hardware evaluation and custom hybrid markers and probe attachment 116
35 Example wire phantom segmentation 118
36 Boxplots of dimensional and angular distortion compared per tracker for the low-cost hardware evaluation 120
37 Side-by-side comparison of typical reconstruction results obtained during the low-cost hardware evaluation 123
38 Side-by-side comparison of typical reconstruction results obtained during the low-cost hardware evaluation 124

List of Tables

1 Initial misregistration offsets for gesture control experiment 68

2 Difficulty levels for MARIN evaluation experiment 91

3 Mean dimensional distortion per axis for each combination of hardware in the low-cost hardware evaluation 121

4 Mean angular distortion per axis for each combination of hardware in the low-cost hardware evaluation 121

5 TRE after registration for each combination of hardware 123

7 Summary of thesis findings 131

List of Abbreviations

2D Two Dimensional

3D Three Dimensional

AD Angular Distortion

ANOVA Analysis of Variance

AR Augmented Reality

AVF Arteriovenous Fistula

AVM Arteriovenous Malformation

CNS Central Nervous System

CPU Central Processing Unit

CSF Cerebrospinal Fluid

CT Computed Tomography

CTA Computed Tomography Angiography

DBS Deep Brain Stimulation

DD Dimensional Distortion

ENT Ear, Nose and Throat

FEM Finite Element Method

FPS Frames Per Second

GLSL OpenGL Shading Language

GPS Global Positioning System

GPU Graphics Processing Unit

HMD Head Mounted Display

HSD Honestly Significant Differences

ICP Iterative Closest Point

iCT Intraoperative Computed Tomography

IGNS Image-Guided Neurosurgery

IGS Image-Guided Surgery

iMRI Intraoperative Magnetic Resonance Imaging

iUS Intraoperative Ultrasound Imaging

LMIC Low-income and Middle-income Countries

LTS Long Term Support

MAD Median Absolute Deviation

MP Mega Pixel (10^6 Pixels)

MRI Magnetic Resonance Imaging

NASA National Aeronautics and Space Administration

NTP Network Time Protocol

OR Operating Room

OS Operating System

PCI Peripheral Component Interconnect

PLA Polylactic Acid

RGB Red, Green, Blue

RMS Root Mean Square

ROI Region of Interest

SD Standard Deviation

SUS System Usability Scale

TLX Task Load Index

TRE Target Registration Error

UDP User Datagram Protocol

US Ultrasound Imaging

VGA Video Graphics Array

VR Virtual Reality

WHO World Health Organization

WLAN Wireless Local Area Network

WPA Wi-Fi Protected Access

CHAPTER 1

Introduction

Surgically treatable disorders and diseases affecting the brain and central nervous system (CNS) have a significant impact on health worldwide. The World Health Organization (WHO) estimates that worldwide there is currently ∼308k new cases of cancer affecting the CNS per year and this number is expected to grow to about ∼435k by 2040 [54]. Other disorders affecting the CNS have a high prevalence and impact too. For instance, cerebrovascular diseases, traumatic brain injury and epilepsy are expected to affect nearly 77 million people, 69 million and more than 50 million respectively by 2030 [1, 21] with stroke being the third leading cause of death in most industrialized countries, after coronary heart disease and cancer. Surgical interventions play an important role in the treatment of these varying CNS disorders. Indeed, it is estimated that 13.8 million patients per year require neurosurgical

interventions [20].

In order to surgically treat CNS disorders, clinicians first diagnose the disease and make surgical plans using preoperative images, such as magnetic resonance images (MRI) or computed tomography (CT). Understanding the individual patient anatomy and having access to the spatial location and extent of a lesion (*e.g.,* a tumour, arteriovenous malformation (AVM), etc.) is crucial to the success of a surgical procedure. Providing surgeons with this type of information in the operating room (OR) has been one of the driving forces behind the development of image-guided surgery (IGS) systems.

While IGS systems have been used in many types of surgical procedures, neurosurgery has been of particular interest. Due to the unique requirements of neurosurgery, many surgical advances were originally tailored for it. Neurosurgery is the only type of surgery for which the organ operated on is almost entirely enclosed in bone. Further, it is a critical structure where it is paramount during interventions to minimize damage to eloquent healthy tissue. Owing to this, image-guided neurosurgery (IGNS) systems or neuronavigation systems (in the rest of the thesis, both IGNS and neuronavigation systems will be used interchangeably) are now used routinely at most centres in developed countries.

IGNS systems work by enabling real-time localization of tools and patient within the OR environment, which in turn enables mapping of preoperative images relative to the patient. A typical surgical scene where a neuronavigation system is used is shown in Figure 1. Akin to the use of a GPS system, which uses satellites to localize and visualize a vehicle on a map of a city, in IGNS surgical tools are tracked and visualized on the preoperative images or plans of the patient, thus guiding the surgeon to the lesion while avoiding damaging critical structures.

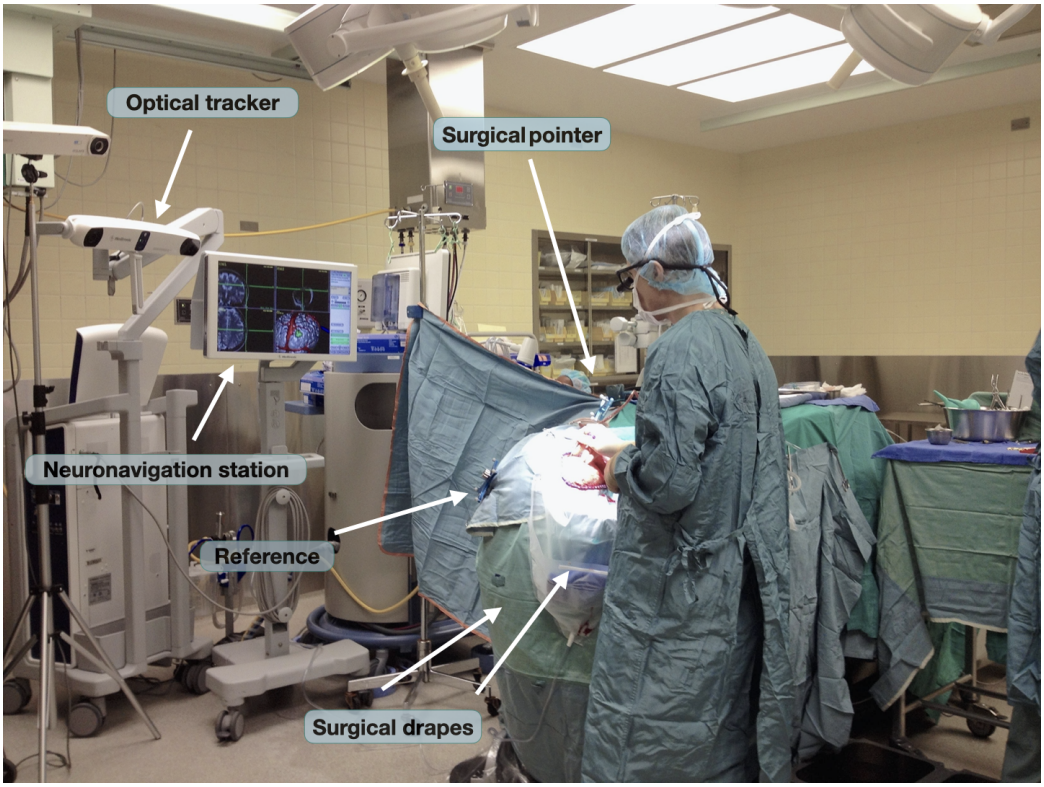

Figure 1: In IGNS the surgeon uses a tracked pointer to relate structures on the patient to those on the preoperative images, shown on the workstation monitor. This pointer and/or surgical tools are localized and tracked by an optical tracking system relative to a reference which is rigidly attached to the patient.

Image-guided neurosurgery systems have enabled neurosurgeons to more accurately and effectively treat CNS disorders. Current commercially available IGNS systems can achieve sufficient accuracy to provide surgeons with meaningful anatomical information during procedures enabling smaller craniotomies (*i.e.,* removal of part of the skull to access the brain) and better informed intraoperative decisions thereby improving surgical outcomes. In tumour resection cases for instance, use of navigation has been shown to lead to significantly lower residual tumour post-resection and reduced neurological deterioration, leading to better quality of life and increased long term survival [81, 131, 112]. IGNS systems have also shown their usefulness in various other neurosurgical applications as well, such as electrode insertion for deep

brain stimulation [133] and neurovascular procedures [73].

1.1 Motivation

While IGNS systems have shown benefits in neurosurgical care, they do have some limitations. The first one of these problems is that of providing intuitive visualization to the user. These systems provide guidance information, but this information is displayed on a monitor, either in orthogonal slices or on a three dimensional (3D) model, which means that the user still needs to translate the information back to the physical patient. This mental mapping of the guidance information to the patient is not always trivial, which makes it error prone, especially for more intricate anatomy. Another shortcoming of current IGNS systems is that the mapping from the patient to the preoperative images that is computed at the beginning of surgery loses accuracy as the procedure progresses. This accuracy loss has multiple causes, but the most prominent of them is the movement and deformation of the brain, or brain shift. This degradation in accuracy causes the guidance information provided to the surgeon to become less and less reliable over the course of the surgery, to the point where it may no longer be usable. A third problem of current systems is their limited interactivity. Once sterile, the surgeon cannot interact with the system themselves, they need to rely on a technician to operate the system for them by following their verbal commands. This interaction by proxy takes longer, leading to both prolonged procedures, and is more error prone. Finally, IGNS systems are expensive, severely limiting their use outside of economically developed countries. Currently, a commercial neuronavigation system, which consists of a workstation loaded with specialized software, tracking hardware and trackable tools, can cost between 250 thousand and 700 thousand US dollars (USD) [72]. This puts these systems out of reach of most hospitals in low- and middle-income countries (LMIC). Of the previously mentioned nearly fourteen

million patients needing neurosurgical interventions every year, it is estimated that 80% arise in LMIC. To fulfill this demand, there is a need for not only a growing neurosurgical workforce, but also access to affordable tools. The question of devising and developing novel neurosurgical guidance systems that are precise and usable, but also low-cost is thus timely, relevant and has potential for a significant impact on global health.

Technological developments in mobile devices, depth and other low-cost sensors can address some of the aforementioned shortcomings and further the state-of-the-art of surgical practice through new hardware integration. New tools enable new visualization and interaction paradigms to be designed and developed allowing for intuitive guidance leading to more streamlined surgical procedures. A significant potential impact of integrating new devices and sensors is the possibility to simplify guidance systems and lower their cost, thus enabling access to them to a much wider audience, potentially saving and improving quality of life post-surgery of up to millions of patients in developing economies.

This thesis explores the integration of mobile devices (*e.g.,* smartphones and tablets) and low-cost hardware alternatives into neurosurgical workflows. The effects of the developed novel visualization and interaction methods are tested and compared against the current state-of-the-art. Thus, the focus of the research presented in this dissertation is to mitigate the shortcomings of current IGNS systems and democratize access to these technologies.

1.2 Objectives and Contributions

The work presented in this dissertation aims at assessing the potential for new technology integration into neurosurgical processes. More specifically, we aim to

assess the use of mobile devices and other low-cost sensors in the context of neurosurgical guidance. The main contributions of this dissertation are:

1. A comparison of *in situ* AR display, *ex situ* AR display and traditional navigation with regards to cognitive load and attention shifting in neurosurgical task execution.

2. The design and development of a state-of-the-art neuronavigation platform implemented on a mobile device, followed by a thorough assessment of its performance and usability.

3. The design and development of a manual registration correction technique implemented on a mobile device using gestures.

4. The assessment of mobile device video images for intra-operatively determining vascular hemodynamics and development of a prototype interaction method to provide hemodynamic information intraoperatively.

5. The evaluation of low-cost tracking alternatives for 3D freehand ultrasound reconstruction in the context of neurosurgery.

This work evaluates the use of mobile and low-cost technologies in the medical domain and as such is of a multi-disciplinary nature. Some of the contributions of this thesis therefore belong to biomedical engineering as well as software engineering. Although, questions at the core of this thesis pertain to evaluating if and how mobile devices and other hardware could be used for neurosurgical guidance applications. To answer these questions, we focus on evaluating computer interaction paradigms in the context of image-guided neurosurgery.

To ensure clinical applicability of the developed methods, all methods were evaluated in either a clinical setting when possible or in a laboratory setting with

input from clinicians and experts. This helps in ensuring that developed methods would be true to their aims and maximize their potential clinical impact.

1.3 Outline

The rest of the thesis is organized as follows. In Chapter 2, an overview of image-guided surgery is given. In Chapter 3, we describe our study comparing *in situ* AR, *ex situ* AR and traditional navigation. Next, in Chapter 4, the development of a gesture-based registration correction method is presented and in Chapter 5, an implementation of a complete neurosurgical guidance system on a mobile device is presented, as well as a thorough evaluation of its performance and usability. As an addendum to Chapter 5, a method to interactively show hemodynamic information intraoperatively is presented. In Chapter 6 a study evaluating the influence of hardware cost on 3D freehand reconstruction and assessment of a low-cost tracking alternative is given. Lastly, we conclude and present avenues of future work in Chapter 7.

CHAPTER 2

Background

This chapter begins with a historical overview of neurosurgical guidance and continues to introduce the components of image-guided neurosurgery (IGNS) or neuronavigation systems. The shortcomings of these systems are then discussed, as well as potential solutions to address these shortcomings. This will lay the foundations for the rest of the thesis, where some of these questions will be explored and potential solutions tested and investigated.

2.1 Image-guided Neurosurgery

Since the proposed application is in the neurosurgical domain, this dissertation will focus on neurosurgery, even though many of the techniques presented below have also been applied to other types of surgical procedures.

2.1.1 History

As summarized by Galloway and Peters in two historical reviews [34, 35] on the development of image-guidance systems, the advances in image-guided surgical procedures have very closely followed the advances in imaging techniques. Indeed, the first reported case of image-guided surgery dates from 1896, when a needle was removed from a woman's hand with the help of a preoperative radiograph. This operation took place only eight days after the announcement of the discovery of X-Rays by Röntgen [109]. A few years later, Horsley and Clarke [50] created what could be considered the first stereotactic guidance system. Stereotactic guidance refers to the family of methods making use of precision tools and techniques to locate internal structures within a patient in order to approach them surgically.

Horsley and Clarke were the first to introduce many concepts which are now considered fundamental to surgery, including the notion of physical- and image-space, perpendicular cutting planes and brain atlases. Although, their most important contribution in the context of this review is the stereotactic frame (see Figure 2) they devised to chart physical space. The frame, which was placed on a subject's head using anatomical landmarks such as the auditory canal and inferior orbital rim, was used to place an electrode at a precise location inside a monkey's brain. Although the system was initially intended to be used for research on monkeys, and not treatment on humans, the main ideas were reused for neurosurgery in the 1940s.

The first stereotactic system devised for human therapeutic use was proposed by Spiegel *et al.* in 1947 [115]. Instead of using external landmarks as before, information was extracted from pneumoencephalograms, *i.e.,* x-ray imaging of the brain where the cerebrospinal fluid (CSF) is first drained and replaced with air to improve contrasts. In this method, a plaster on the patient's head with a metal ring attached was cast, then once dried, a pneumoencephalogram was acquired, enabling a wire or cannula

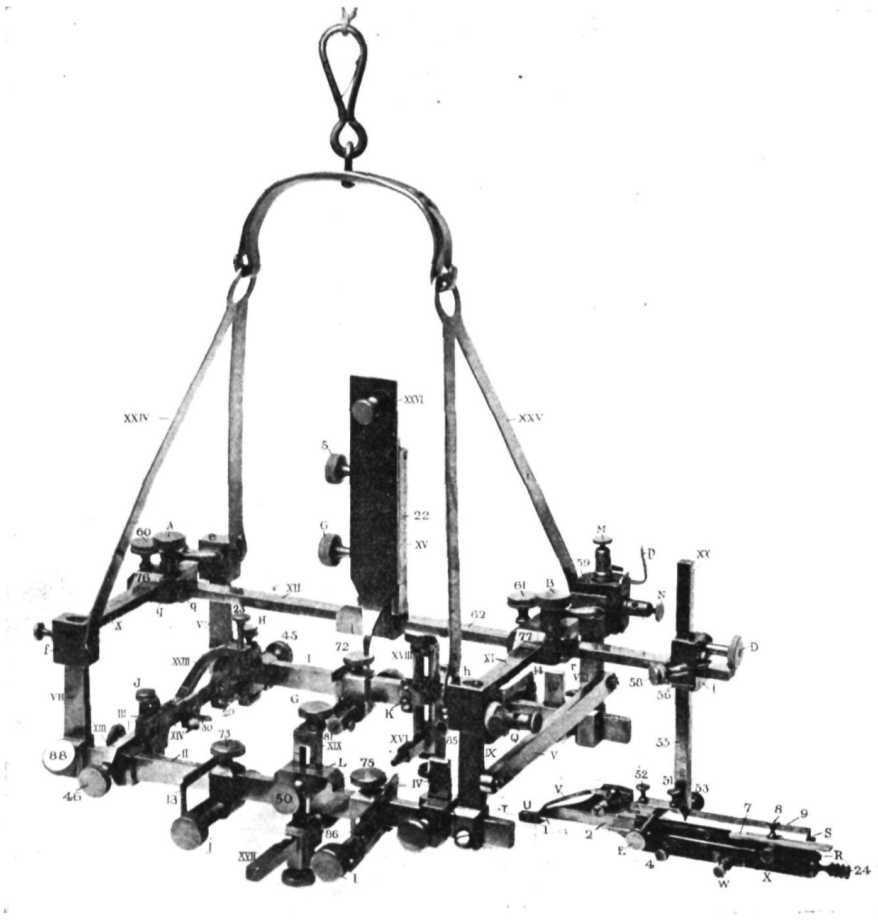

FIG. 7.

Side Elevation of Clarke's instrument.

Figure 2: Horsley and Clarke's stereotactic frame [50].

to be inserted at a precise location. The system was tested for medial thalamotomy, showing promising results, but it was also hypothesized that it could be used in other applications such as the production of pallidal lesions to treat involuntary movements or electrocoagulation of the Gasserian ganglion to treat trigeminal neuralgia, among others. Stereotactic guidance was quickly studied by other groups [119, 70, 103] working on different diseases that all requiring precise localization of structures within the brain. The rise in popularity of stereotaxy in those years was mostly due to new techniques such as ventriculography and pneumoencephalography, which

provided surgeons with patient-specific information to help them map structures. The possibility to accurately place a narrow-gauge needle using guidance images led to other applications, such as placing a radiation source [4, 90] or a chemotherapy agent [77] directly within a tumour and to guide a biopsy and permit a much smaller craniotomy [3].

The next big breakthrough happened in 1973, when Hounsfield [52] showed how it was possible to build a three-dimensional volume or computed tomography (CT) of a given anatomy out of multiple two-dimensional X-ray slices. The system was soon adopted and tested in a clinical setting by Bergström and Greitz [5]. For the first time in history, clinicians had access to three-dimensional preoperative images during a surgical procedure. This led to the development, by Roberts *et al.* [106] and other groups [129, 66, 32], of the first navigation systems where the positions of tools was tracked in physical-space and the corresponding position in image-space was displayed. The systems developed by these teams varied in implementation and hardware: Roberts *et al.* and Friets *et al.* [106, 32] used ultrasonic tracking while Watanabe *et al.* and Kosugi *et al.* [129, 66] used a tracked arm. Roberts *et al.* and Friets *et al.* projected the images to be viewed in a neurosurgical microscope, while Watanabe *et al.* and Kosugi *et al.* showed images on a computer display. This family of systems was first referred to as "frameless stereotaxy", since contrary to previous systems they didn't require the patient's head to be fixed in a frame. This was also the first use in the literature of the expression "image-guided intervention", which is now widely used.

These "image-guided intervention" systems, while differing in implementation and hardware had commonalities. They consisted of an image acquisition step, three-dimensional spatial tracking of the tool in physical-space, a registration process to map the physical-space into image-space and finally a display of the rendered

localization information. The paper from Roberts *et al.* [106], where images were projected into the microscope, was also among the first to use an augmented reality (AR) display for image-guidance (see Figure 3).

Augmented reality, according to the definition of Milgram *et al.* [84], is "augmenting natural feedback to the operator with simulated cues". In other words and in our context of interest, AR is the merging of additional virtual information (*e.g.,* anatomical models) to the live surgical scene. The concept of AR represented a paradigm shift compared to traditional systems. As well as mapping the real-world, *i.e.,* the pointer tip location, onto the images, here the images themselves were also mapped onto the real-world.

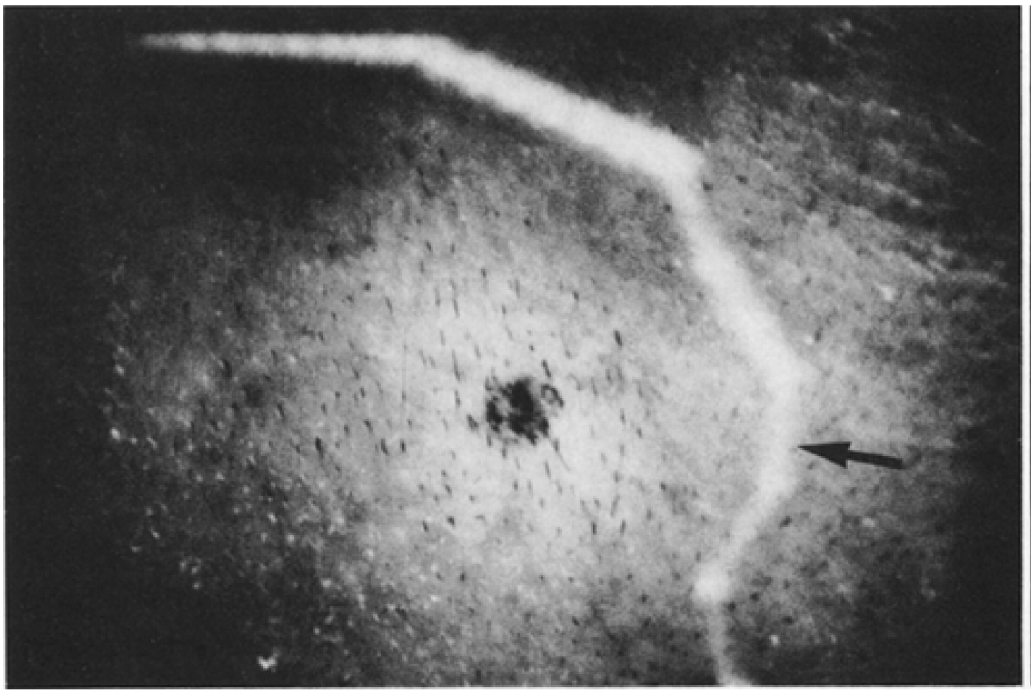

Figure 3: Roberts *et al*'s [106] microscope augmented reality view. (The tumour outline is represented by the white line pointed by the black arrow.)

The idea of augmented reality was studied further during the next decade by Roberts' team at Dartmouth, and soon other teams followed [45, 65, 27]. The

developed research systems used either the surgical microscope as the display device on which they blended real and virtual images or a projector displaying virtual information directly onto the patient. The surgical microscope was the easiest and most natural candidate for displaying the augmented reality view: it was already present in the operating room (OR), the surgeon was already using it for the most part of the operation and it was relatively straightforward to add a video stream to the live view already presented. While augmented reality in the microscope is still an active area of research, in the recent years, new display technologies have arisen. Different display types and the way they have been and could be used in the OR will be further examined in subsection 2.3.1.

As was shown in this brief overview of the development and advancement of image-guidance frameworks, the number of applications for such systems is vast. Although, as will be shown in section 2.2 and section 2.3, there are limitations to current systems. The interested reader can find a thorough overview of current applications of image-guidance systems in [56]. In the next section we examine current navigation systems in more detail, with the hardware and software that is most often used.

2.1.2 IGNS System Components

IGNS systems such as those described above are no longer merely used for research but are also commercially available systems. Complete systems, like the ones marketed by Medtronic® [76] and Brainlab® [9], are now currently used in most modern operating theatres. These systems have three main requirements: a tracking system that allows for the localization of the surgical tools and the patient, a registration that correlates the patient images to the actual patient in the operating room, and a workstation that provides the visualization of the tracked tools with respect to the patient anatomy and surgical plans. Figure 4 shows a typical surgical installation.

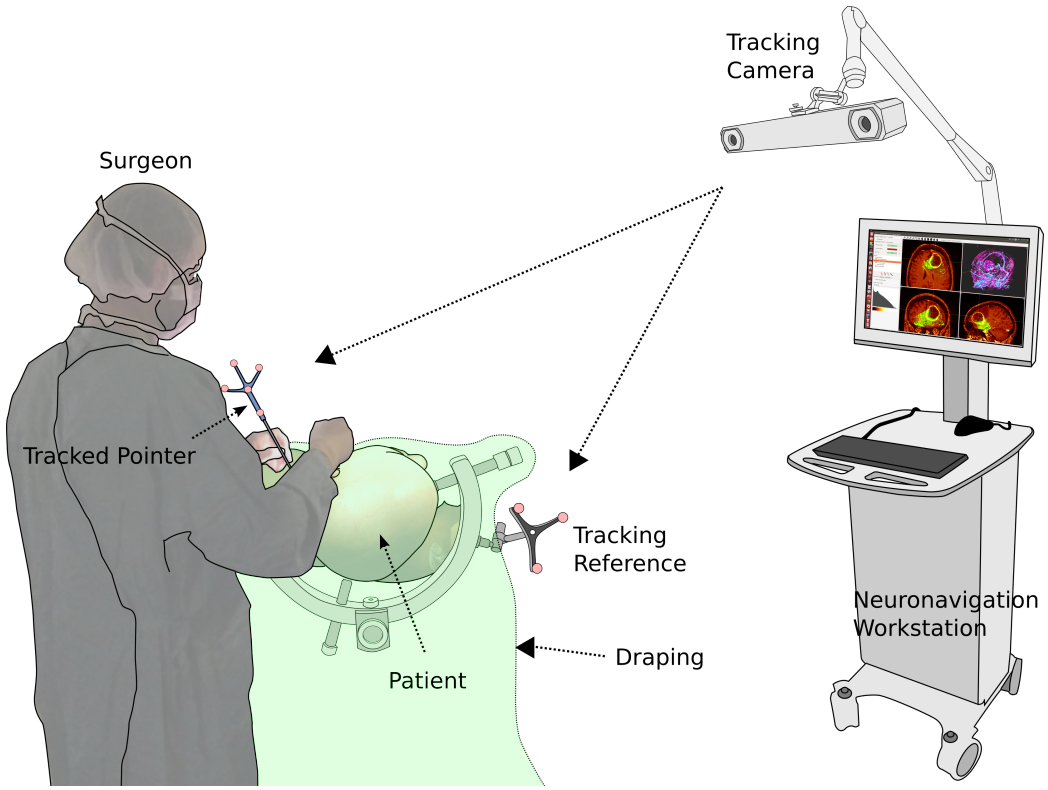

Figure 4: IGNS systems have three main requirements: a tracking system that allows for the localization of the surgical tools and patient, a registration that correlates the patient images to the actual patient in the operating room, and a workstation that provides the visualization of the tracked tools with respect to the patient anatomy and surgical plans. Figure courtesy of Simon Drouin.

Tracking

Tracking systems are used to localize objects (*e.g.,* surgical tools, patient) in 3D space. To be useful, a tracking system must be able to provide both the position and orientation of the tracked objects. This makes the problem of tracking six-dimensional, *i.e.,* translation in Cartesian x, y and z coordinates and rotation along each axis, or yaw, pitch and roll. In the context of IGNS, four types of tracking systems have been used: mechanical, acoustic, optical and electromagnetic.

Mechanical Tracking uses a mechanical arm equipped with angular sensors in its joints to determine the position of a tracked tool. The tool is attached to the end of

the arm and the position is computed compounding the angles and segment lengths for all segments of the arm. Although these systems have very high accuracy, their mobility is very limited and usage can be disruptive to normal surgical workflows. While in the context of open neurosurgery this limited mobility can be problematic, for robotic surgery this is not an issue as trajectories can be planned in advance, such as in the case of electrode placement for deep brain stimulation (DBS).

Acoustic Tracking uses emitters and three or more microphones. The difference in reception time of the sound produced by the transmitter on a given tool enables triangulation of its 3D position. While acoustic systems were studied in the beginning of IGNS development, their use has almost disappeared now due to limited accuracy, limited number of measurements per time interval and sensitivity to changes in room temperature affecting the speed of sound [30].

Electromagnetic Tracking systems use a generator to produce a magnetic field within the operating room. Tracked objects are equipped with coils that interfere with the magnetic field and the position of the coil within can be calculated based on the interference. This class of systems has the main advantage that it doesn't require line of sight between the emitter and receivers, contrary to both acoustic and optical tracking. However, it also has significant drawbacks: tracked objects need to be wired, which can be inconvenient in the sterile context of the surgical field, and any metallic or ferromagnetic object introduced in the magnetic field will disrupt the field and alter the reading.

Optical Tracking systems, which are the most widely used in neurosurgery [37], rely on stereophotogrammetry. Reflective spheres are attached to tracked objects, usually three of four, and the reflection of infrared lighting on these spheres is captured

by a tracker equipped with a set of two infrared cameras. The difference in position of the spheres in the two images enables computation of positions of the spheres in space and thus in turn, knowing the particular arrangement of the spheres on the object, the position and orientation of the tracked object can be computed.

Registration

Registration involves spatially transforming two or more data sets into one coordinate system. In the case of neuronavigation systems, it means transforming the coordinates of the patient anatomy in physical-space so that it correlates with the preoperative images (*e.g.,* CT or MRI), in image-space. This requirement of navigation systems appeared when evolving from stereotactic frame-based systems to frameless systems. With frame-based systems, the spatial relationship between the tool and the relevant anatomical structure could be determined since both the mapping from the tool to the frame and the mapping from the frame to the anatomy were known, and the compounded transform could easily be computed. With frameless systems, this mapping is not readily available, therefore there is a need for an explicit registration procedure. This patient-to-image registration is typically done once, prior to beginning the surgical procedure, after the patient is anesthetized and the patient's head is immobilized using a head clamp (such as the Mayfield® clamp). Seeing as the coordinate system in physical space must be fixed relative to the patient to remain usable throughout surgery, instead of using the tracking system itself as the reference frame, a marker is attached directly to the head clamp and serves as the reference for navigation. Registration estimates a rigid transform relating the patient's head position to that reference position.

The two methods used for registration in the OR are point-based and surface-based. Point-based methods rely on the ability to accurately pick corresponding

points in the two spaces – points in image-space are chosen on the preoperative scans and points in physical-space are chosen on the patient using a tracked pointer. From the known transform of each of these points, the global transform can be computed. In order to compute a global transformation matrix three pairs of points are needed. However, the robustness of the method can be increased by picking more pairs in order to minimize the residual error of the resulting transform. Common points selected for point-based registration in neuronavigation are the bridge of the nose, the external and internal canthi of the eye and the tragus and tragus valley of the ear, since they are easily recognizable and can be accurately selected on the preoperative images and on the patient's head. A schematic of this registration procedure is shown in Figure 5.

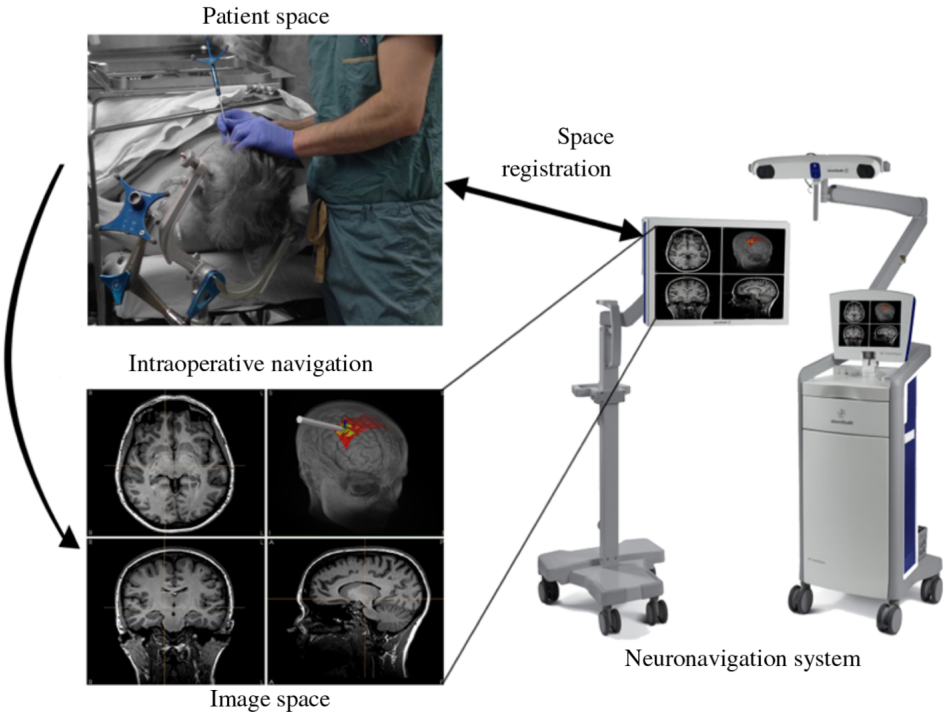

Figure 5: Schematic of the registration procedure. Eight corresponding landmarks on the head and face are selected both on the patient using a tracked pointer and on the preoperative images to create a correspondence between the two spaces. Figure taken from [38].

Surface-based registration relies on being able to extract a surface in both

physical and image space and minimize the distance between the two extracted surfaces, usually using an iterative method, such as the iterative closest point (ICP) algorithm [7]. The surface in physical-space is created by tracing the pointer across the skin on the head of the patient. In image-space the skin surface is extracted from the preoperative CT or MRI.

Neuronavigation Workstation

This last component consists of a computer with specialized IGNS software. The IGNS software enables tracking, *i.e.,* the information from the stereo infrared tracking cameras are imported here and used to extract the 3D position and orientation of the patient and all tracked tools. Algorithms for registration and the various transforms computations that enable merging of physical-space and image-space are available. The workstation also allows for visualizing the result of all these computations and provides guidance to the surgeon throughout the procedure, see Figure 6.

2.2 Accuracy of IGNS Systems and Brain Shift

In spite of the remarkable advances in the field, current IGNS systems have limitations in certain scenarios. The two most prominent of those shortcomings are (1) the loss of accuracy that happens throughout the procedure and (2) the difficulty in presenting guidance in a clear, understandable and interactive fashion to the operating surgeon. In this section, we focus on the first of these shortcomings and describe the main causes for accuracy loss and some ways to correct for it. The next section then describes problems associated with interaction and visualization.

Brain shift or the displacement and soft-tissue deformation of the brain that occurs during surgery, is the main source of loss of accuracy in IGNS systems. The error caused by brain shift typically ranges from 1 to 20 mm [39], which is much higher

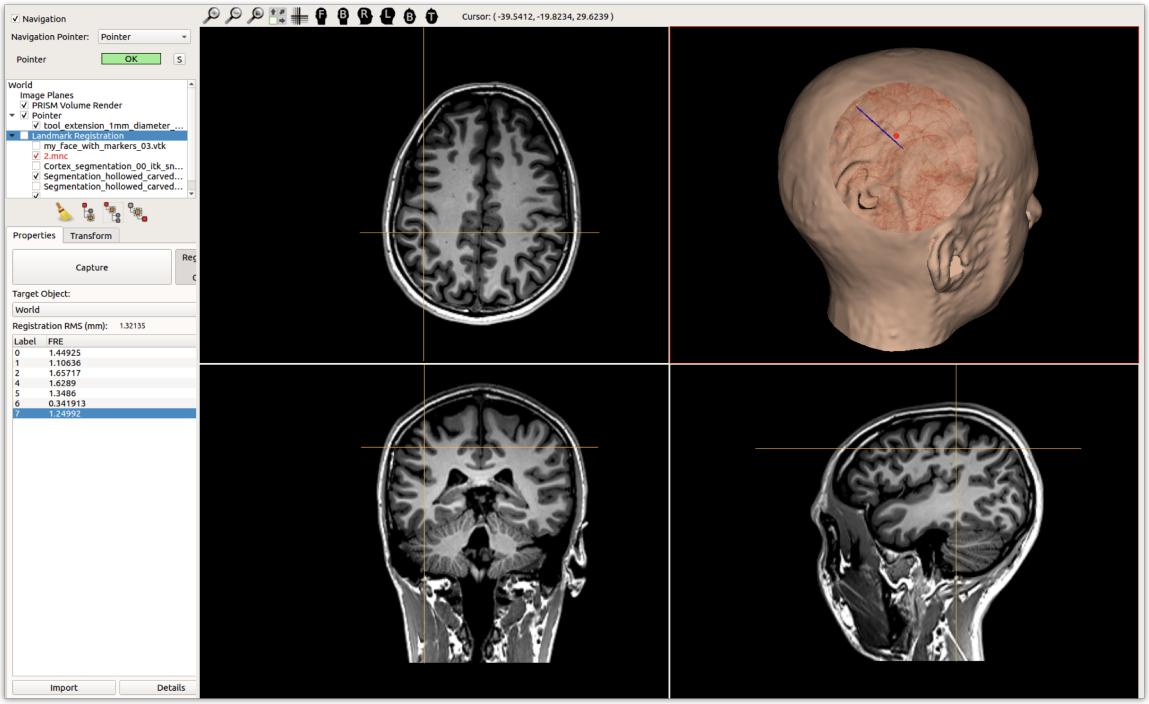

Figure 6: Screenshot of a typical IGNS navigational interface. More specifically this screenshot depicts the Intraoperative Brain Imaging System (IBIS) [25]. Slices from the preoperative scan are shown, as well as 3D models of different anatomies, in this case the skin surface, cortex surface, as well as a target of interest displayed as a red dot. In the left pane are options to enable users to interact with the view and change behaviour of the system.

than errors arising from technical inaccuracies, such as inaccuracies in the tracking, distortion in the preoperative images or registration errors. The error introduced by brain shift is sometimes large enough that surgeons no longer rely on navigation information after the first few stages of surgery [39]. An example of brain shift estimation acquired through intraoperative MRI acquisition is shown in Figure 7.

Brain shift is the result of a complex process involving multiple factors related to biological, surgical, and physical causes. In the first reports on brain shift, the cause for brain shift was attributed to gravity, but it is now known that there are many other contributors. While the phenomenon of brain shift has been known for more than 30 years [59], studies isolating the individual contribution of all different

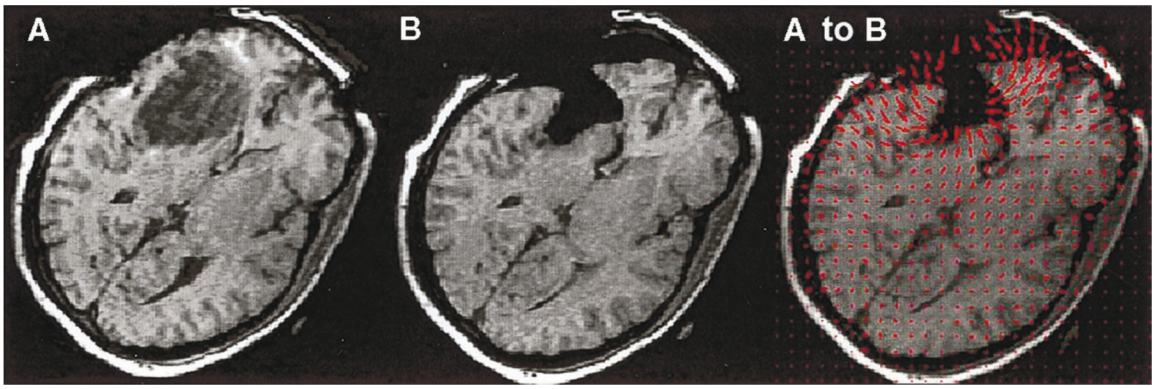

Figure 7: Successive MRI acquisitions taken during surgery. Left: Image acquired after the dura mater opening; Center: Image acquired after tumour resection. Right: Estimated deformation field (red arrows) representing the shift observed after tumour resection. Figure taken from [91].

factors are lacking. In this section, we first survey the causes of brain shift and their potential contributions, then look at methods developed to compensate for it.

2.2.1 Causes of Brain Shift

Brain shift causes are multiple but can be categorized in three main classes: physical, surgical and biological. Here we summarize the main known causes for brain shift, but the interested reader can see [116] for a more detailed review and [126] for a review of neuronavigation error sources.

Surgical and Physical Factors: The transform relating the patient to the head clamp is assumed to stay constant throughout surgery. However, this often is not the case. The registration procedure is done prior to sterilization and during sterilization, the reference is removed and replaced with a sterile reference, which can introduce displacements relative to the initial position. Next, the clamp itself does not fully rigidly hold the patient's head, it allows for small movements, which can be larger if strong forces are applied to the head. Such large forces are applied during burr hole hand drilling or hand sawing of the cranium for instance. Use of other surgical

equipment can also have the effect of moving the head with respect to the frame. For instance, skin retractors, which are used to keep the skin flap away from the surgical field and improve bone exposure, apply a lateral force of 10-15 N [39]. That lateral force will be compounded if multiple retractors are used, so if three retractors are applied, which is a typical usage, then the resulting lateral force applied on the head would be around 30-45 N. This would result in several millimetres of translation. After registration, draping is placed around the surgical field to ensure it stays sterile. The drapes are usually attached to different parts of the surgical field, including the clamp, again applying a force on it. The drapes are pulled in different directions when installed. During surgery, the draping will often absorb blood and cerebrospinal fluid (CSF), increasing their weight and resulting in additional downward forces. So, even the simplest assumption of registration, a fixed relationship between the head and the tracking reference, isn't usually respected.

Other surgical factors include tissue loss, fluid loss and hyperosmotic drug intake. Tissue loss is due to resection of pathological tissue; it causes unsupported surrounding tissue to sag due to the force of gravity. Fluid loss refers to drainage of CSF. For many types of procedures, CSF is purposefully drained to allow for brain relaxation and reduction of intracranial pressure. CSF loss induces a deformation mainly in the direction of gravity due to loss of homeostatic neutral buoyancy, causing the brain to deform under it's own weight. Intracranial pressure is also commonly reduced by injecting Mannitol, an hyperosmotic drug, which decreases brain volume. Mannitol intake thus considerably changes the shape of the brain and therefore invalidates the correcteness of preoperative images as a navigation basis for the state of the brain after intake. In addition to the surgical interactions described, patient positioning also affects accuracy of the navigation system. The orientation of the patient's head intraoperatively is almost never the same as the orientation of the

head during the preoperative scan. Patient positioning, as detailed by Schnaudigel in [111], has an impact on the position of the brain within the skull and even it's shape. Deformations of up to 1.7 mm were observed between MRI scans acquired with patients' heads placed in different orientations. It still isn't clear how much of this deformation is due to anisotropy in the magnetic field causing image distortion and how much by the brain's movement and deformation within the skull. Further studies would be necessary to assess this, but in both cases, precision of the navigation would be equally affected. These effects, contrary to the ones presented in the last paragraph are non-linear effects and thus harder to account for and compensate.

Biological Factors: In addition to the factors mentioned above, it was also shown by Dorward *et al.* [22] that tumour type, in the case of tumour surgery, has an effect on brain shift. Doward *et al.* hypothesized that the brain tends to herniate out of the defect depending on the pathology type. Indeed, the authors found significantly greater shifts at the point in surgery where the surgeon had reached the depth of the tumour for meningiomas compared to gliomas and significantly less shift in skull base cases than other groups. Although, on the contrary, Ohue *et al.* [94] found significantly smaller shifts for meningiomas than gliomas. Replication studies with larger sample sizes are necessary, however these studies would indicate that tumour type could perhaps influence brain shift. Finally another factor related to lesion type is presence or absence of œdema, a build up of fluid, which also influences the direction of the shift.

Although it is widely accepted that gravity is the largest cause of brain shift, as we have described above, other causes are numerous. The causes are not only many but all interrelated and hard to isolate. Studies assessing the individual contributions of each factors are lacking, but will be necessary for compensation techniques to improve. The more methods are developed to compensate for larger contributors, the

more important it will become to understand lesser contributions as well.

2.2.2 Brain Shift Compensation

With a better understanding and knowledge of the causes of brain shift, methods can be devised to try and account for it. The methods proposed thus far can be categorized into two main classes: methods where new images are acquired intraoperatively to reregister preoperative images to the patient and methods that use biomechanical and predictive modelling to estimate brain shift and then warp the intraoperative images to reflect the shift. Both approaches have advantages and disadvantages. Intraoperative-imaging methods are more versatile as they use updated images and thus can compensate for all effects of brain shift at once, even those caused by unforseen events or for not well-understood reasons. However, these methods require intraoperative aquisition of new images, which may be expensive (*e.g.,* in the case of intraoperative MRI (iMRI)), disrupt the normal workflow and prolong the procedure (*e.g.,* for each iMRI acquisition, 45 mins or more is added to the procedure). On the other hand, model-based methods require no or very little intraoperative information and intervention, and can be very precise for some compensations. Although they cannot compensate for changes caused by surgical factors since these methods rely on precomputation of deformation and surgical factors cannot be known *a priori*. They also can't compensate, for the same reason, for shift caused by unexpected events or due to causes that are hard to isolate.

2.2.3 Intraoperative Imaging for Brain Shift Correction

Intraoperative image acquisition for brain shift compensation is typically done using either intraoperative MRI (iMRI) or ultrasound (iUS), although intraoperative CT (iCT) and surface acquisition through 3D scanning have also been studied. iMRI

has the advantage that acquired images have higher contrasts than iUS and the registration is easier to compute and more precise since intraoperative images will usually be from the same modality as preoperative images. iMRI also has very important shortcomings: it is very expensive and it requires a specifically designed OR, equipped with compatible tools, free of any metal or ferromagnetic material, as well as specially trained staff. Further, making a single acquisition takes between 45 minutes to an hour thus prolonging the procedure. The long acquisition time often prohibits making multiple acquisitions at different points in surgery. Finally, due to the prohibitive cost, very few centres are equipped with iMRI scanners, therefore we focus on ultrasound based intraoperative registration.

The first article presenting use of iUS clinically in brain tumour resection procedures was published in 2002 by Unsgård *et al.* [122]. Ultrasound was shown to be useful in other procedures such as AVM or aneurysm clipping as well, as described in Unsgård *et al.*'s subsequent review on ultrasound use for neurosurgery [123]. Ultrasound, contrary to MRI, is cheap, portable, does not have any special requirements and allows for quick image acquisition, which makes it possible to do multiple acquisitions during a single procedure if needed. Although, registration of iUS with preoperative MRI is harder to accomplish since both modalities have different contrasts and the image quality of ultrasound is lower in terms of resolution, signal to noise ratio and acoustic resonance artefacts. Several methods have been proposed to register B-mode iUS images with preoperative MRI including: mutual information [55], gradient orientation alignment [92], cross correlation [105] and linear correlation of linear combinations [33]. In addition to the B-mode ultrasound images, Doppler acquisition can also be used for registration. Indeed, Doppler US captures movement and thus can be used to visualize blood vessels. The intraoperatively acquired vessel tree can then be mapped to a similar tree extracted

from a preoperative angiogram. This method was first proposed by Reinertsen *et al.* in 2007 [99] and further validated clinically [102, 100]. The same idea has then been further exploited by Morin *et al.* [88] in a more complex algorithm using a combination of intraoperative image acquisition registration and biomechanical modelling.

As well as accounting for brain shift, US can show residual tumour tissue. Unsgård showed that iUS acquisition helped find residual tumourous tissue in 53% of the cases whose resection would have otherwise been deemed complete. They reported taking between three to six iUS aquisitions, which is much more than what would typically be possible with iMRI. More recent studies [40, 127] show that even though ultrasound's precision is lower than iMRI, it is still sufficient to enable surgeons to identify tumour remnants in many cases.

In addition to iUS imaging, some teams have studied the possibility of using cortical surface extraction to compensate for brain shift. Various methods have been explored for cortical surface extraction, *e.g.,* stereo-camera surface reconstruction [117] or laser-range surface scanning [82]. These methods showed promising results, enabling near real-time estimation of the surface movement. While it is clear that these methods can be a valuable source of information [12, 67, 136], it has also been shown that accuracy of the predicted brain shift improves when using more than only the visible surface [132].

2.3 Visualization and Interaction with Guidance Information

A second shortcoming of IGNS systems is the difficulty to present spatial information efficiently and clearly to surgeons. With traditional navigation systems, navigation information is presented on the workstation which is outside of the sterile region in

the OR and thus requires the surgeon to look away from the patient and surgical field for guidance. Surgeons are then required to mentally map 2D patient-specific images (*e.g.,* MRI/CT), or 3D rendered anatomical models to the actual 3D anatomy of the patient. This task is not trivial, can be time consuming and may be prone to error [61]. Furthermore, it causes a disconnect between where the surgeon is working and where they are looking for guidance. This constant shifting of attention can be detrimental to the surgeons' ability to complete a given task, as will be further explained in Chapter 3. Disruption to the user's focus has been shown to be detrimental to both cognitive and motor tasks. Pearson *et al.* [95] and Postle *et al.* [98] showed that focus shifts have a significant negative effect on performance of tasks involving spatial working memory. Furthermore, the effects of distractions on the completion of a motor task worsens with the complexity of the task in question [44].

Another problem associated with the current systems' information presentation is the difficulty of showing the dimensionality and conveying spatial understanding of the anatomical data. When 3D structures are presented on a 2D screen, the sense of depth is often lost or at least severely hindered.

Then, because of sterility constraints, the surgeon also can't interact with the system themselves. Interacting with the presented view can enable gaining a better understanding of the 3D structure of the anatomy and spatial relationship between structures. Manipulating it in such a way as to make the mapping to the patient easier also has potential to reduce errors. Since such interactions can't be done by the surgeon themselves, a technician does those manipulations according to the surgeon's needs. This interaction by proxy is not ideal though, as it adds unnecessary complexity and potentially lengthens the procedure.

Both of these shortcomings can be addressed using specific visualization and interaction techniques and/or AR display technologies, as some of these AR devices

offer means for the user to interact with the displayed information, even in sterile conditions. Several AR hardware and display means have been proposed in the context of neurosurgery to address these issues. We describe these related works in the following section.

2.3.1 Augmented Reality Image-guided Neurosurgery

In augmented reality neuronavigation, preoperative images or virtual models of the patient anatomy are merged with the real surgical field of view to guide the surgeon (see Figure 8). AR has been shown to be useful in many aspects of neurosurgery: Cabrilo *et al.* [11] showed that AR permitted minimizing dissection and exposition in the context of aneurysm clipping; Kersten-Oertel *et al.* [61] showed that AR enabled tailoring craniotomies, localizing vessels of interest, and easily distinguishing between feeding and draining vessels in the context of arteriovenous malformation (AVM) procedures; and Gerard *et al.* [38] showed that using AR creates the possibility of tailoring resection corridors to minimize invasiveness, especially for deep-seated tumours.

Figure 8: Left: computer-generated image of vasculature, generated from the preoperative scan; Center: image taken from a tripod-mounted camera; Right: displayed AR view of the surgical scene.

An Overview of AR Systems for IGNS

The first augmented operating microscope prototype, proposed in 1986, was aimed at cranial surgery [106]. The authors used an acoustic tracking system, a miniature cathode ray tube and a beam splitter to superimpose a tumour boundary in the microscope's optics. They found that their system provided easily visualized contours while not diminishing the quality of the conventional operative field image. The first camera-based AR systems for neurosurgical procedures were proposed by Gleason *et al.*and Gildenberg *et al.*in 1994 [42, 41]. Gildenberg *et al.* [41] attached an endoscope to the stereotactic frame and displayed a segmented tumour and vessels on top of the camera image on a computer monitor. The registration was done using a stereotactic frame. While Gleason *et al.* [42] used a video mixer to blend streams from a video camera and 3D computer reconstruction of the segmented anatomy. They used landmarks in the video image to align the reconstruction with the live view. In 1995, the first augmented stereo microscope prototype was proposed by Edwards *et al.* [26]. The work showed that displaying structures in stereo added a sense of depth, which was previously lacking. The first use of AR for endovascular procedures, done in 1988 by Masutani *et al.*, displayed vessels extracted from preoperative angiograms [74]. Lastly, the first endoscopic AR application was developed in 2002 by Kawamata *et al.* [58]. They tested their system in endonasal transphenoidal procedures.

In terms of surgical applications, of the 33 papers presenting recent AR platforms studied in the review by Guha *et al.* [47], 48% were applied to tumour resection, 27% to neurovascular surgeries and 21% for spinal procedures. Although there is a predominance of neurooncology applications, different types of surgical procedures have been explored. Whatever the target application is however, many display technologies can be employed. The interested reader is referred to [47] for a table summarizing the neurosurgical applications in which AR has been studied. This table

shows that more than 50% of the proposed systems employ a screen display, meaning that the information is displayed outside of the sterile surgical field. Therefore, in order to access navigation information the surgeon has to look away or shift their attention between the surgical field and the navigation. This constant shifting of attention can be detrimental to the surgeons' ability to complete a given task, as will be shown in Chapter 3. However, many display technologies enable *in situ* display in the context of surgery: augmented optics on the microscope, HMDs, projectors and mobile devices. Each of these have advantages and drawbacks making them more or less advantageous in different scenarios and applications.

As mentioned in the historical review above, the microscope was an obvious choice in the early days of neurosurgical AR navigation and is still an option currently being explored by research teams. This AR setup uses special optics to blend the augmentation image directly in the microscope oculars. Its biggest advantage consists in that it is already present in the OR and thus doesn't require any additional footprint in an already crowded room. It is also used by many surgeons throughout most of the procedure, making it a negligible disruption to the normal workflow. Although, in the typical procedure, it is only used once the *dura mater* is open, making it irrelevant at the planning stage and the beginning of the procedure, when the craniotomy, or surgical opening of the skull, is performed. It also isn't used by all surgeons or for all types of procedures. This display thus doesn't cover all potential use cases.

In the past few years, HMDs have been the center of much interest in the AR neuronavigational research community. HMD systems can either be optical see-through or camera-based. In the former case the device uses a semitransparent display to show the augmentation, through which the real world underneath can still be seen. In the later case, the real world is captured with a camera and the blending of the augmentation and live view is rather done computationally and shown on a

screen in front of the user's eyes. Like microscopes, their added footprint is negligible but an advantage over microscopes is that they can be used in all stages of the procedures, including during planning. Although, they still have disadvantages that limit a mainstream adoption in ORs. Recent studies by Carbone and Condino [18, 14] even suggest that in its current form, see-through devices are unsuitable for surgical guidance. They are afflicted by the vergence-accomodation conflict. This problem arises from the fact that the virtual content presented is at a fixed optical distance from the user's eyes. This distance is not the same as the content's perceived distance, which causes the conflict, as the user's visual system adjusts to the real world scene. Virtual structures thus appear blurry and prolonged use of the system can cause eye strain, headaches and nausea. This problem affects all currently available headsets and while it is currently being investigated, it still doesn't have a satisfactory solution. Additionally, HMDs require adjustments when fitted to the user's head in order to work properly and accurately. This can add complexity and disrupt the normal workflow as, once the surgeon is sterile, they can't fit it on their head or remove it themselves, requiring help from a technician. This implies some setup time for every use and also means that their is a risk of a maladjustment, which could compromise the accuracy of the presented AR view. Additionally, Condino's research also hints that the eye-to-display calibration that the tested device implemented may lead to distortion, leading to an inaccurate AR view. Finally, currently commercially available devices are also still bulky and uncomfortable when used for extended periods. For all of these reasons current HMDs are poorly suited for many surgical tasks.

Projector-based AR neuronavigation systems use a projector to present the augmentation directly on the surgical scene. They can be used at every step of a procedure like HMDs, but do not cause discomfort if used for long periods unlike

them. Although, they require an additional setup and have an OR footprint. More importantly, they require line of sight between the projector and the surgical scene which, in the context of a busy operating room, can be a significant drawback and add logistical constraints.

Finally, mobile AR uses a mobile device, *e.g.,* a phone or tablet, to capture the real world from the back-side camera, computationally generate the AR view through texture blending in the graphics shader and displaying it on the screen of the device. An example of such a system is showed in Figure 9. As will be further shown in Chapter 3, Chapter 4 and Chapter 5, mobile devices have several advantages over all previously mentioned methods:

- They are small and wireless, which makes them easy to move around. This allows exploring the anatomy from different perspectives, providing a good understanding of the three-dimensional structures of objects through the parallax effect. This portability also easily allows the surgeon to explain the approach and anatomy to residents and students in the operating room. Additionally, the device can be moved out of the way at any time and is therefore less encumbering and more suited to the surgical environment in comparison to HMDs and projectors.

- They can be draped in a sterile bag, allowing for continuous use throughout a procedure. This also means they can be stowed away and brought back in a matter of seconds, making them minimally disruptive to the surgical workflow.

- They can be hand-held and moved around, as mentioned, but they can also be clamped to the bed for hands-free continuous guidance, accommodating for different operating practices and methods.

- They offer an easy means to not only display information but also interact with it using the touchscreen, allowing more and easier interactivity than projector-based or microscope-based AR systems to the surgeon without exterior intervention. Touchscreen interactions can also be done through sterile draping, allowing for interaction at every stage of surgery.

- They are equipped with additional embedded sensors (*e.g.,* accelerometer, gyroscope, depth sensor) that haven't been explored in this context yet, but could potentially provide useful guidance information. This point is further discussed in Chapter 7.

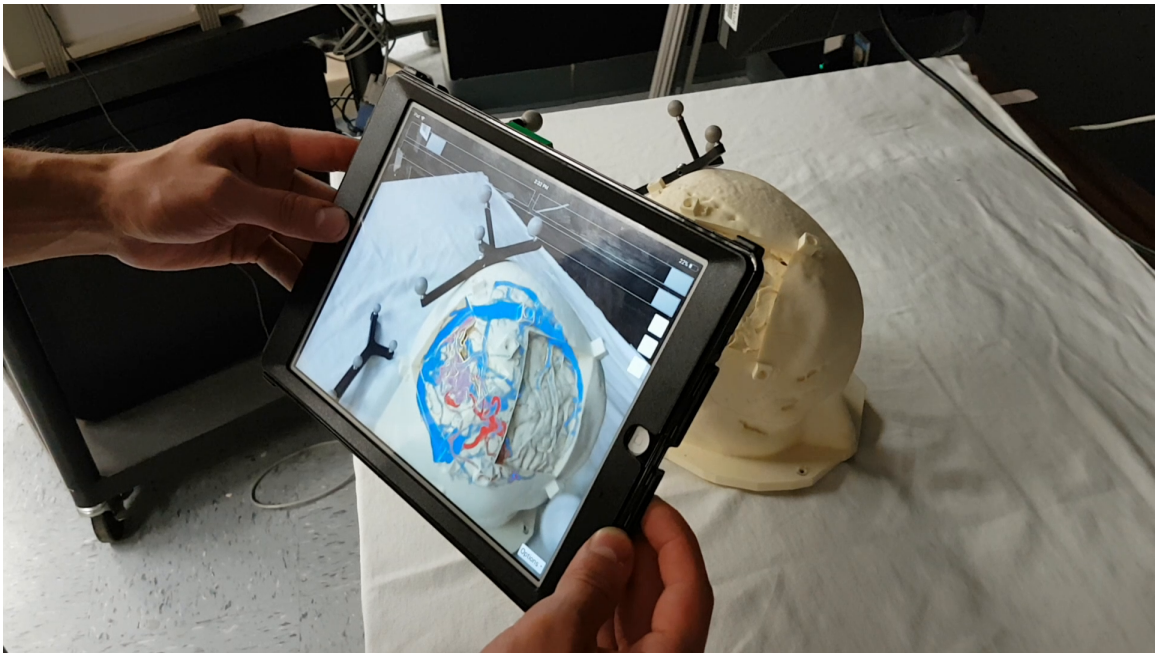

Figure 9: Our developed mobile AR system displaying CTA acquired vessels (virtual anatomical data) over a phantom head (the real world as captured by the iPad).

Despite the advantages mentioned above, this technology has had limited study in the context of neurosurgery. The camera-based mobile AR prototypes proposed thus far are those listed in Table 10. These prototypes are few and all are lacking either in terms of usability, performance or in thoroughness of system assessment.

From this table, it appears clearly that mobile AR has not been fully explored for neurosurgical AR guidance. A more thorough description of the previously proposed systems is presented in Chapter 5, Section 5.2.

Reference	Tracking & registration method	Visualization method	Clinical validation	Error measurement
Eftekhar *et al.* [29]	No tracking, manual registration (visually align images)	Blend preoperative MR slices with live view	5 external ventricular drain insertions	- In plane measurement in pixels assuming no rotation - Between 8-14mm difference with Medtronic navigation system
Hou *et al.* [51]	No tracking, manual registration (visually align images)	Blend preoperative MR slices with pre-captured sagittal picture	15 + 20 intracranial lesions	- Measurement on CT slices (15 cases) assuming no rotation - Uncertainty on measurement greater than observed values - Offset vector magnitude from Medtronic navigation system (20 cases), 1.5-14.2 mm
Watanabe *et al.* [128]	6-camera VICON system with landmark registration (3 landmarks)	3D rendered image of the virtual head overlaid on live view	6 tumour resections	- In plane distance computed in lab, to be around 1 mm (methodology not explained) - No measures made in clinic.
Deng *et al.* [19]	Optical tracking using Polaris Spectra with landmark registration	3D rendered structures overlaid on live view	2 + 3 tumour resections	- In plane measurement of distances between manually selected landmarks in pixels 1-8.6 pixels
Chang *et al.* [16]	Kalman filtered localization from fiduciary marker	Live view of the Solid models of structures overlaid on live view of the patient	None	- RMS error in targeting within 2.2 mm (methodology not explained)
Bieck et al. [8]	Optical tracking using a Polaris Vicra, registration unspecified	3D rendered surfaces	None	None

Figure 10: Table summarizing existing mobile AR neuronavigation applications.

The remainder of this dissertation is organized as follows. In Chapter 3, the problem of visualization is addressed and the potential for AR use in neurosurgical tasks is evaluated. In Chapter 4, a solution for brain shift correction is developed. This solution runs on a mobile device and can be interacted with by the surgeon, also addressing the interactivity and visualization problems simultaneously. In Chapter 5, the work from Chapter 3 and 4 is expanded and consolidated. The more complete system presented in that chapter further explores solutions to both interactivity and visualization issues. As an addendum to the chapter, a specific workflow is developed for neurovascular navigation, focusing on these problems in that context. In Chapter 6, low cost tracking alternatives are evaluated for brain shift correction.

CHAPTER 3

Quantifying attention shifts in augmented reality image-guided neurosurgery

Preface

Using a mobile device for presenting guidance information opens up the possibility for *in situ* augmented reality. In this case, a phone or tablet can be used by the surgeon or surgical staff within the surgical field and thus the surgeon does not have to look away for guidance to the neuronavigation system. Thus, using a tablet has the possibility of reducing attention shifts which have been shown to negatively impact motor and cognitive tasks. In this chapter, we evaluated the impact of image-guidance location in a typical neurosurgical task, tumour delineation for

cranitomy planning. Specifically, we evaluated the impact of three intraoperative set-ups: standard navigation (traditional neuronaviation set-up), desktop AR (AR visualization on the neuronavigation monitor) and mobile AR (AR visualization *in situ*) on attention shifts and usability. The results of our study confirmed potential clinical relevance of mobile device for AR capture and display. Furthermore, in addition to showing potential clinical relevance, results also informed the design decisions of the methods later described in Chapter 4 and Chapter 5. This paper was presented at the Augmented Environments for Computer Assisted Interventions (AE-CAI) workshop and published in a special issue of Healthcare Technology Letters.

Abstract

Image-guided surgery (IGS) has allowed for more minimally invasive procedures where smaller incisions are used, leading to better patient outcomes, a reduced risk of infection, less pain, shorter hospital stays and faster recoveries. One drawback that has emerged with image-guided surgery techniques, is that the surgeon must shift their attention from the patient to the monitor for guidance. However, it has been shown that there are negative affects of shifting attention on both cognitive and motor tasks. Augmented reality, which merges the real world surgical scene with preoperative virtual patient images and plans has been proposed as a solution to this drawback. In this work we studied the impact of two different types of augmented reality IGS set-ups (Mobile AR and Desktop AR) and traditional navigation on attention shifts for the specific task of craniotomy planning. We found a significant difference in terms of time to perform the task and attention shifts between traditional navigation but no significant differences between the two different AR set-ups. However, with Mobile AR users thought the system was easier to use and their performance was better. These results suggest that regardless of where the AR visualization is shown to the surgeon, for certain tasks AR can reduce the amount of attention shifts, leading perhaps to more streamlined and focused procedures.

3.1 Introduction

Augmented reality (AR) is increasingly being studied in image-guided surgery (IGS) for its potential to improve intraoperative surgical planning, simplify anatomical localization, and guide the surgeon in their tasks. In AR IGS, preoperative patient models are merged with the surgical field of view, allowing the surgeon to understand the mapping between the surgical scene and the preoperative plans and images and to see the anatomy of interest below the surface of the patient. This may facilitate

decision making in the operating room (OR) and reduce attention shifts from the IGS system to the patient, allowing for more minimally invasive and quicker procedures.

Numerous technical solutions have been proposed to present AR views in IGS. These include the use of tablets, projectors, surgical microscopes, half-silvered mirrors, head-mounted displays (HMDs), or the use of the monitor of the IGS system itself [63]. These different solutions can be categorized as either presenting the AR visualization within the field of view of the surgeon, *i.e., in situ* (*e.g.,* via tablets, HMDs, the microscope or a projector) or outside the surgical sterile field on the IGS system itself. Whereas, the main advantage of the former, is that the surgeon does not have to look away from the surgical scene, the disadvantage is that additional hardware is needed, with the exception of the surgical microscope. Although the surgical microscope can present the AR view to the surgeon without the use of additional hardware it is not used for all surgical steps or by all surgeons. For example, it is not used during craniotomy planning or by surgeons who prefer to use surgical loupes throughout a case. Conversely, the advantage of using the IGS system to display the AR view, is that no additional hardware is needed in an already cluttered and busy OR. The disadvantage of presenting the AR view on the IGS monitor may be that the surgeon may need to shift their attention back and forth between the IGS system (where they are looking for guidance) and the patient (where they are working). To the best of our knowledge no previous work has looked at the impact of different AR solutions on attention shift for specific tasks in IGS. Yet, evaluating how different technologies compare to Desktop AR is an important task, which will help determine which technologies are most appropriate in the OR.

In this paper, we present a mobile-based (*e.g.,* smartphone/tablet) AR IGS system and compare it to (1) visualization of the AR view on the monitor of the IGS system and (2) traditional IGS navigation. We do this for the specific task of outlining the

extent of a tumour on the skull (Figure 11). Tumour localization and delineation is done during craniotomy planning in neurosurgery in order to determine the location, size and shape of the bone flap to be removed to access the brain. This task does not require high accuracy in terms of outlining the tumour contour but rather it is important to localize the extent of the tumour. Therefore, in testing these three methods we specifically focus on the time to delineate a tumour and the number of attention shifts from patient to screen, that are required to do this.

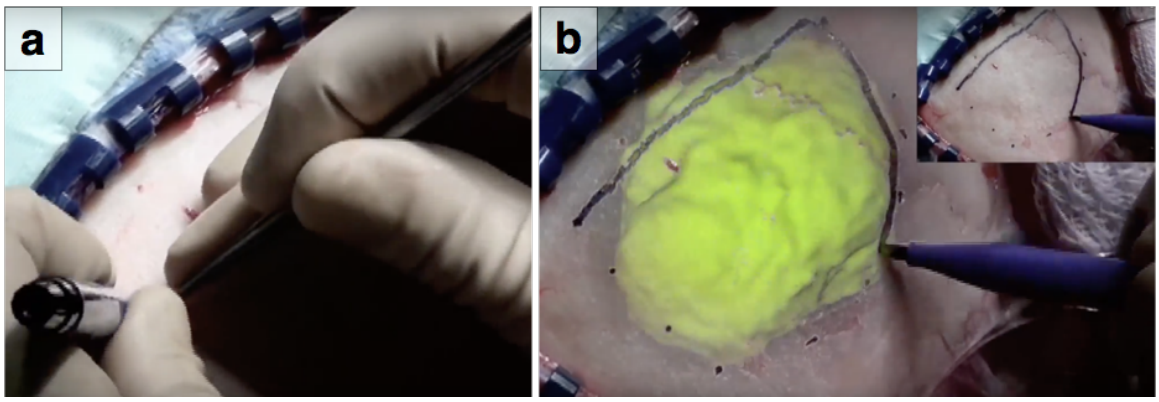

Figure 11: (a) A surgeon uses a pointer in his right hand to locate the boundary of the tumour and draws dots with his left hand at different locations. (b) AR visualization would allow the surgeon to see the tumour merged with the real surgical scene and can use that to draw the extent of the tumour. (b) *Inlay:* with traditional neuronavigation this surgeon has drawn dots using guidance and then connects the dots to create the contour of the tumour.

3.2 Related Work

In the following section, we first give a review of related work focusing on of the use of AR in image-guided neurosurgery (IGNS). Second, we explore previous work on attention shifts in surgery.

3.2.1 AR in Image-guided Neurosurgery

The first neurosurgical AR system was proposed in the early 1990s by Kikinis *et al.* [64]. In their work they combined 3D segmented virtual objects (*e.g.,* tumours) from preoperative patient images with live video images of the patient. For ear, nose and throat(ENT) and neurosurgery, Edwards *et al.* [28] developed MAGI (Microscope-Assisted Guided Intervention), a system that allowed for stereo projection of virtual images into the microscope. Varioscope AR was a custom built head-mounted operating microscope for neurosurgery that allowed for virtual objects to be presented to the viewer using VGA displays. The Zeiss OPMI Pentero's microscope and its Multivision function (AR visualization) was used by Cabrilo *et al.* [11] for augmented reality in neurovascular surgery. One of the findings of this work was that surgeons believed that AR visualization enabled a more tailored surgical approach that involved determining the craniotomy. Kersten-Oertel *et al.*, used AR visualization in a number of neurovascular [61] and tumour surgery cases [62]. In both studies the authors found that AR visualization (presented on the monitor of the IGNS system) could facilitate tailoring the size and shape of the craniotomy.

Over, the last several years mobile devices have been increasingly used to display augmented reality views in order to ease and speed-up several tasks in surgery. Mobasheri *et al.* [86] presented a review of the different tasks for which mobile devices can be used, including diagnostics, telemedicine, operative navigation and planning, training, etc. To the best of our knowledge there has not been research that has examined using mobile AR specifically for craniotomy planning. However, Deng *et al.* [19] and Watanabe *et al.* [128] have built mobile neuronavigation AR systems, which they test in surgery including craniotomy planning. Then, Bieck *et al.* [8] introduced an iPad-based system aimed at neurosurgery. Hou *et al.* [51] also built an iPhone-based system to project preoperative images of relevant anatomy onto the

scalp. Some prototypes of mobile AR for surgery have also been tested in other contexts, for example in nephrolithotomy by Müller *et al.* [89]. We have expanded on this previous work in AR IGNS by looking at how different AR display methods, specifically in the surgical field via mobile device versus on IGNS monitor, may impact the surgeon in terms of attention.

For more information as to the use of augmented reality in image-guided surgery the reader is referred to [63], and augmented reality in neurosurgery specifically [47, 118, 78].

3.2.2 Attention Shifts in Surgery

As summarized by Wachs [125], attention shifts have negative effects on surgical tasks. In general, attention shifts can deteriorate performance. The work of Graydon *et al.* [44] and Weerdesteyn *et al.* [130] showed how distractions and attention shifts impact various types of cognitive and motor tasks such as counting backwards and avoiding obstacles while walking. Goodell *et al.* [43] showed how surgical tasks in particular are impacted. They observed an increase of 30-40% in the time required to complete a task when a subject was distracted compared to when they were not distracted. In our work, we study the number of attention shifts needed to perform a simple surgical planning task using both augmented reality and traditional navigation. We did not focus on accuracy of the tracings however, previous work has shown that in both a lab and clinical environment, AR guidance is no less accurate than the traditional navigation systems [6, 11].

3.3 Methodology

Our IGS system comprises of a Polaris Tracking System (Northern Digital Technologies, Waterloo, Canada) and the IBIS Neuronav open-source platform for image-guided neurosurgery [25]. IBIS runs on a desktop computer with a i7-3820 3.6 GHz CPU, NVIDIA GTX670 GPU, ASUS PCE-AC55BT (intel 7260 chipset) wireless PCI card and Ubuntu 14.04.5 LTS (with the latest available wireless drivers (iwlwifi 25.30.0.14.0). To extend the functionality from IGS to mobile AR IGS we use a smart phone device (OnePlus One phone with a Qualcomm MSM8974AC Snapdragon 801 chipset, Quad-core 2.5 GHz Krait 400 CPU, Adreno 330 GPU and Android 6.0.1.) outfitted with a passive tracker that is attached to a case to obtain the live view.

The IBIS Neuronav package comes with plug-ins for tracking, patient-to-image registration, camera calibration, and the capability to do augmented reality visualization by capturing a live video stream from a microscope or video camera and merging this with preoperative images on the monitor of the system itself. In our work, we extended the IBIS Neuronav system to allow for augmenting an image, not only on the monitor of the system but on a mobile device that captures the "surgical" field of view. Thus allowing for *in situ* AR visualization.

To make use of IBIS' existing functionality, the mobile device serves merely as a camera and display. The costly computations are handled by the desktop on which IBIS runs. In order to create the AR view, we first calibrate the camera of the mobile phone. Calibration (intrinsic and extrinsic) is done using a modification of Zhang's camera calibration method[137], followed by a second optimization procedure to find the transform from the tracker to optical center of the camera (for more details the reader is referred to [25]). Patient-to-image registration is done using skin landmark registration. For desktop AR, the mobile device captures live video frames and sends them to the desktop using OpenIGTLink [120]. These live frames are then augmented

with virtual objects, in our case the 3D surface of a tumour. For mobile AR, the rendered virtual object is sent using OpenIGTLink to the mobile device on which it is blended with the live video feed using OpenGL (version ES 3.0) and GLSL. The Qt framework (version 5.8) was used to handle the phone's camera and create the AR mobile phone application.

3.4 Experiment

To determine the impact of using AR with *in situ* visualization (Mobile AR) in contrast to AR visualization on the navigation system (Desktop AR) and traditional neuronavigation (Traditional Nav) on attention shifts, twelve subjects (aged 24-41, 3 female and 9 male) working in medical neuroimaging and/or image-guided surgery did a laboratory study. The subjects were graduate students, researchers, engineers and neurosurgery residents. All subjects were familiar with image-guided neurosurgery and craniotomy planning.

The task of the subjects was to draw the contour of a segmented tumour on the surface of the skull of a phantom – a task typically done during craniotomy planning in tumour resections, see Figure 11. Prior to the study the subjects were re-familiarized with the purpose of tumour delineation, craniotomy planning, and augmented reality visualization and were shown the system under each of the conditions (described below). The order in which the different systems were used in the experiment were alternated between subjects and each of the possible condition orders were used an equal number of times to reduce learning bias.

To perform the task, subjects used a permanent marker to draw the tumour outline on a 3D printed phantom that was covered in self adhesive plastic wrap. Each subject delineated four segmented tumours that were mapped to the 3D phantom under each of the three conditions: Mobile AR, Desktop AR, and Traditional Nav.

For Mobile AR the tumour was blended with the camera image on the phone whereas for Desktop AR, the AR visualization was shown on the monitor of the IGNS system. Lastly for Traditional Nav, both the tumour and the head of the patient were rendered in order to give the subject contextual information about the location of the tumour, see Figure 12. For each delineation task, subjects could decide whether to use the surgical pointer and see its location on the IGNS system/or phone with respect to the tumour. However, regardless of if they used it or not, they always held it in their hand throughout the experiment. The set-up of the experiment is shown in Figure 13.

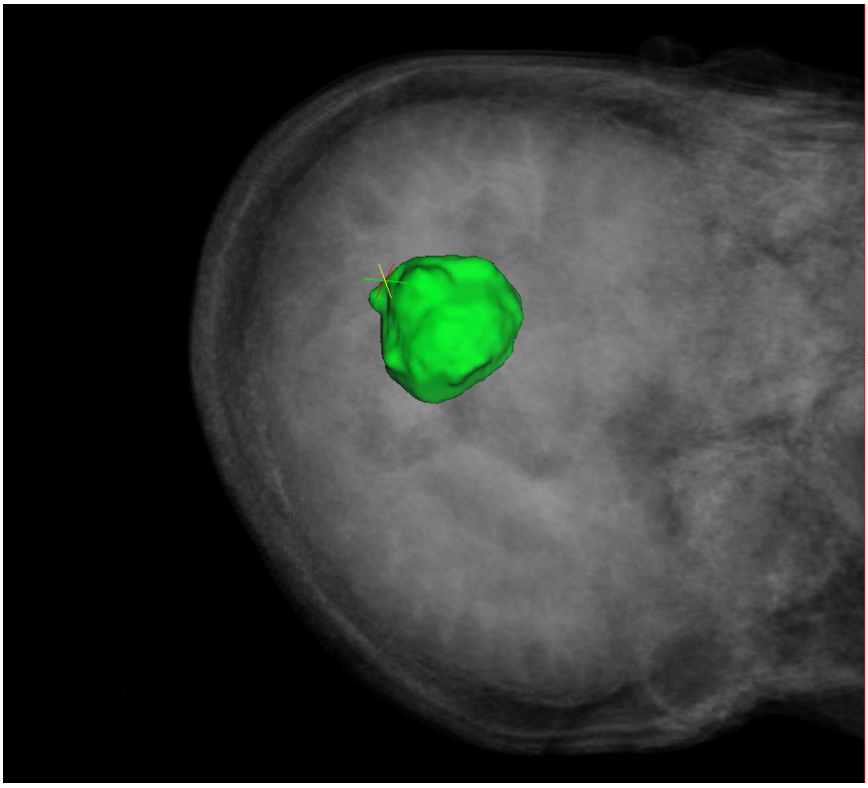

Figure 12: Screen shot of the IGNS monitor view for the Traditional Nav condition: The user has access to the pointer position (colored cross-hair) as well as patient's preoperative scan and the segmented tumour model (green).

Whereas in traditional IGS a surgeon must make use of the surgical pointer to determine the location of it with respect to the virtual anatomy of the patient and surgical plans, this is not necessary when using *in situ* AR as the virtual data is

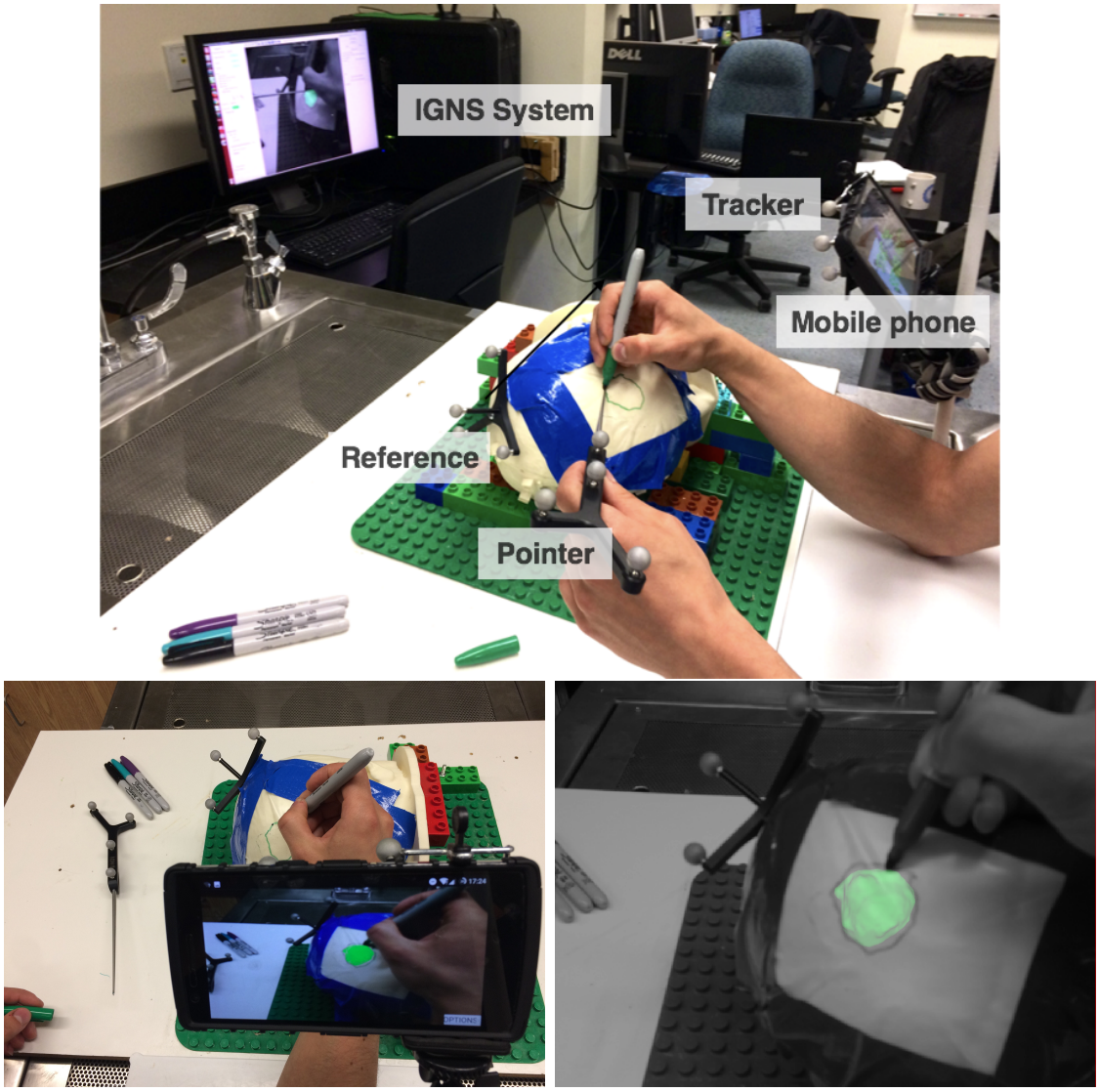

Figure 13: Top: Experimental set-up: The user holds the pointer in one hand and the marker in the other. Depending on the condition he or she looks either at the mobile phone (outfitted with a tracker) for the AR visualizion or on the Desktop for either AR or IGNS navigation. The purpose of the task is to draw the contour of the tumour on the surface of the phantom. Bottom left: Subjects' point of view of the experimental set-up when testing the Mobile AR condition. Bottom right: Screen shot of the Desktop AR view.

visible in the surgical field of view. We therefore, allowed each subject to decide how and whether to make use of the surgical pointer and under which conditions to use it. For the Mobile AR condition, the phone was attached to an "arm" that remained in place throughout the study. Although allowing for these differences

between conditions could potentially lead to confounds in the time taken to delineate the tumour, we believe it allows for the most realistic scenario and one which would mimic how a surgeon would work under the different conditions in the operating room. For example, the surgical pointer would be used with traditional neuronavigation but not necessarily *in situ* AR, making the task take longer under the non-AR condition. Lastly, subjects could also decide whether to use a "connect-the-dots" strategy (mark dots on the phantom at the edges of the tumour and draw a line between them) or simply outline the tumour with the contour.

For each of the tasks we measured both the time to complete the task as well as the number of times the subject switched their attention from the 3D phantom to the IGNS system or mobile phone.

After performing the experiment all subjects filled out a questionnaire[1]. The questionnaire includes the NASA TLX (which pertains to the perceived workload of using the system) [48], as well as a number of other questions about their experience. Furthermore, subjects were asked to provide any additional comments on performing the task under each of the different conditions.

3.5 Results

In terms of the system itself, the pointer calibration was 0.24 mm RMS error, the registration between phantom and virtual models was 1.76 mm RMS error and camera's intrinsic reprojection error was 1.75 mm. For the Mobile AR system, a framerate between 15 and 20 FPS at a resolution of 640×480 was achieved (without compression).

For the experiments, we measured the time it took to delineate the tumour and the when and for how long they looked at the 3D phantom, at the mobile phone or at

[1]The questionnaire can be found in Appendix A

the monitor of the IGNS system. We analyzed the data using an analysis of variance (ANOVA) and post-hoc Tukey honestly significant difference (HSD) tests. The JMP Statistical Software package and Matlab were used. As well as looking at time, and number of attention shifts, we also looked at the ratio between the amount of time the subject looked at one of the screens in comparison to the total time taken to delineate the tumour.

For all the measures, we found that both AR systems were statistically different from the traditional navigation system, but they were not statistically different from one another. Specifically, a one-way repeated measures ANOVA showed that there was a significant effect of AR display type on the total time to delineate the tumour ($F(2, 136), p < 0.0001$). The mean times for tumour delineation were 50.78 ± 24.34 seconds (s) for Traditional Nav, $25.5 \pm 10.95s$ for Desktop AR and $20.6 \pm 8.23s$ for Mobile AR. Post-hoc Tukey HSD tests showed that there was a significant difference between Traditional Nav and Desktop AR ($p < 0.0001$) and Mobile AR ($p < 0.0001$), but that there was no significant difference between Desktop AR and Mobile AR ($p = 0.3134$).

For total number of attention shifts from the phantom to either desktop or mobile screen, a one-way repeated measures ANOVA showed that there was a significant effect of IGNS display type on the number of attention shifts ($F(2, 136), p < 0.0001$). The median number of attention shifts during tumour delineation were 26, with a median absolute deviation (MAD) of 7.5, for Traditional Nav, 4 (MAD=4) for Desktop AR and 1 (MAD=1) for Mobile AR. Post-hoc Tukey HSD tests showed that there was a significant difference between Traditional Nav and Desktop AR ($p < 0.0001$) and Mobile AR ($p < 0.0001$), but that there was no significant difference between Desktop AR and Mobile AR ($p = 0.1075$).

For ratio of total time spent looking at the display over total time taken to

delineate the tumour, a one-way repeated measures ANOVA showed that there was a significant effect of display type on the ratio ($F(2, 136), p < 0.0001$). The mean ratio of time spent looking at the screen over time taken during tumour delineation was 0.60 ± 0.18 for Traditional Nav, 0.91 ± 0.07 for Desktop AR and 0.95 ± 0.05 for Mobile AR. Post-hoc Tukey tests showed that there was a significant difference between Traditional Nav and Desktop AR ($p < 0.0001$) and Mobile AR ($p < 0.0001$), but that there was no significant difference between Desktop AR and Mobile AR ($p = 0.3649$).

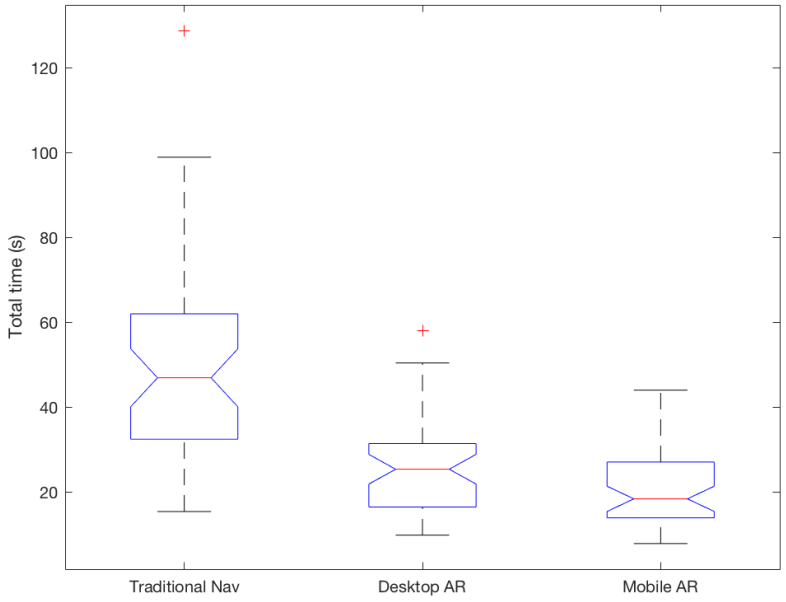

Figure 14: Boxplots of the total times taken per condition (in seconds). The average times to delineate a tumour were 50.78 ± 24.34, 25.5 ± 10.95 and 20.6 ± 8.23 for Traditional Nav, Desktop AR and Mobile AR respectively.

3.5.1 Questionnaire Results

All subjects filled out a questionnaire after performing the study. As mentioned in the last section, we asked the subjects to hold the pointer in all conditions, however,

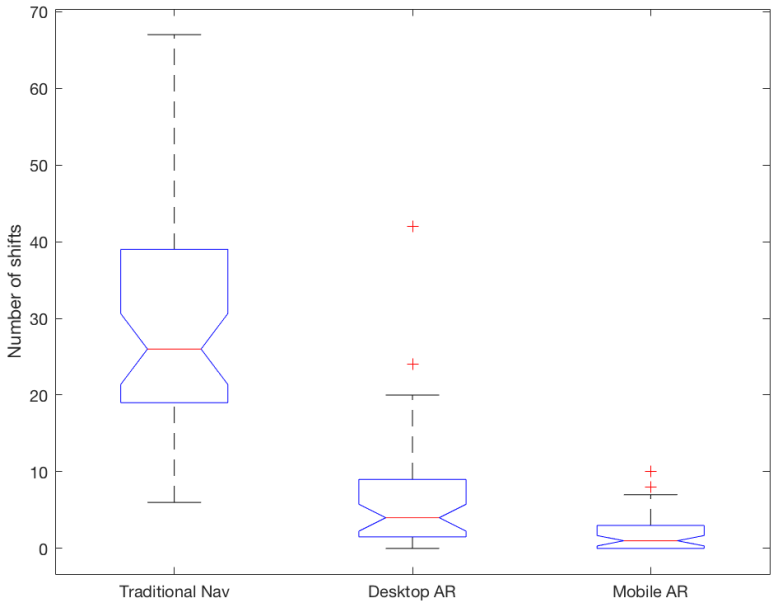

Figure 15: Boxplots of the number of attention shifts per condition. The average number of attention shifts were 27.8 ± 14.00, 6.3 ± 7.79 and 2.3 ± 2.93 for Traditional Nav, Desktop AR and Mobile AR respectively.

according to the post-task questionnaire 58% of the subjects did not use it for Desktop AR and 67% of subjects did not use it in the Mobile AR condition. In terms of user reporting of accuracy, only one subject found that he/she was more accurate with Traditional Nav. All others found AR to be more accurate, specifically 67% found that Mobile AR was the most accurate. Further, all subjects found that one of the two types of AR was most intuitive and comfortable, of those 83% thought that Mobile AR was the most intuitive and 92% thought Mobile AR to be the most comfortable. Overall, 92% of subjects preferred Mobile AR. Finally, the TLX confirms those last findings since, on average, the traditional Nav scored 59 points, the Desktop AR 46 points and Mobile AR 39 points, where a lower score means the perceived cognitive load was less.

As we can see in Figure 17, Mobile AR is perceived as the least demanding system to use overall, followed by Desktop AR. In terms of mental demand, subjects found

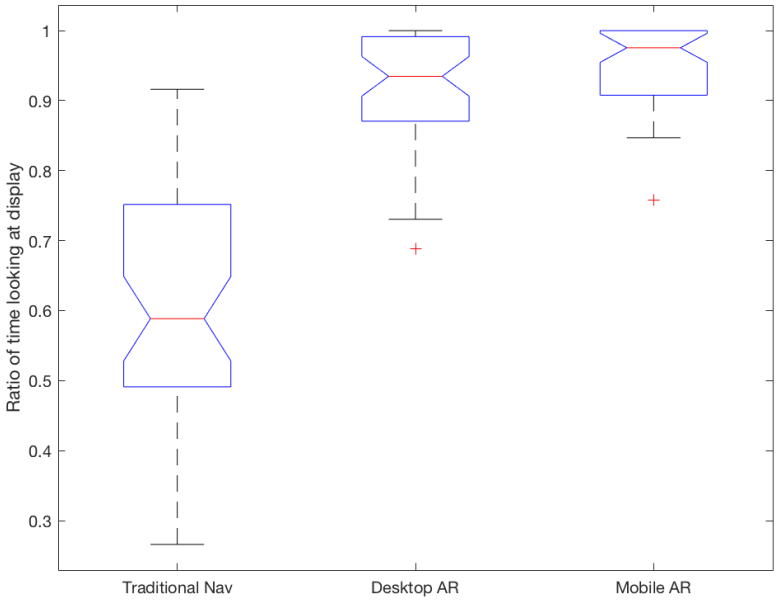

Figure 16: Boxplots of the ratio of time looking at desktop/mobile over total time per condition. The averages were 0.60 ± 0.18, 0.91 ± 0.07 and 0.95 ± 0.05 for Traditional Nav, Desktop AR and Mobile AR respectively.

both AR systems to be equally less demanding than Traditional Nav. Although, in terms of physical demand, temporal demand, effort and frustration, Mobile AR was perceived as less demanding than Desktop AR, which was less demanding than Traditional Nav. In terms of performance, Mobile AR was perceived as best, again followed by Desktop AR.

3.6 Discussion

Our results showed that for both AR guidance methods, the attention of the subject remains almost the whole time (90-95%) on the guidance images. In contrast, for Traditonal Nav the attention is split almost 50-50 between the patient and the monitor. Such attention shifts can be detrimental to the motor task at hand, add time to performing the task, and may increase the cognitive burden of the surgeon.

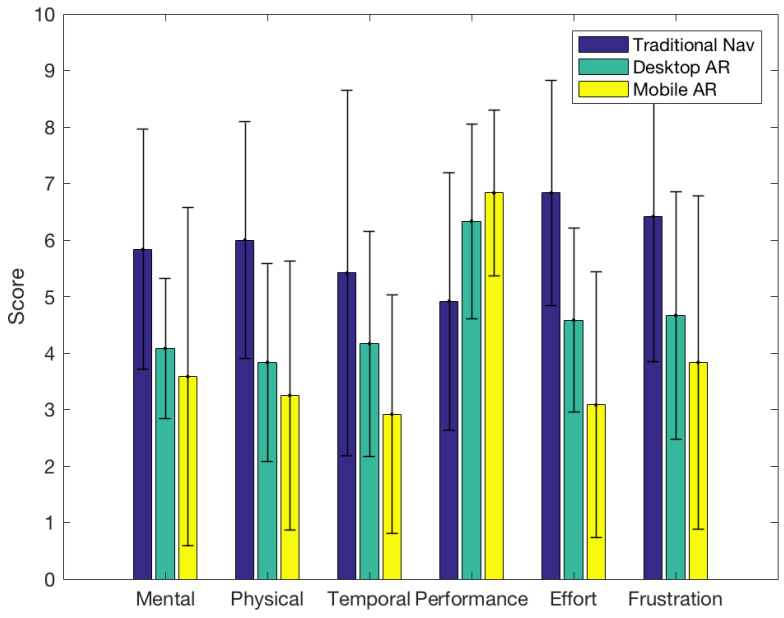

Figure 17: Individual scales of the NASA-TLX for the different conditions, ranging from 1 to 10 (lower is better on all scales except Performance). The results show that for all measures Mobile AR was perceived to be better/easier to use than Desktop AR, which in turns performs better than Traditional Nav.

The ratio of time looking at the screen against total time taken, may also give an estimate of the user's confidence in what he/she is doing. The users shift to look at the patient when they need to confirm that the pointer and the marker are where they expect them to be. Consequently, the higher ratios obtained for both AR systems may indicate that AR gives users more confidence that they are correct with respect to the data presented. The NASA TLX results in terms of perceived performance further seem to support this claim with subjects being most confident in their performance using AR.

Our results also show that the time needed to accomplish the tumour outlining and the number of attention shifts done during the task are significantly lower when using either of the AR systems than when using traditional navigation. Although Mobile AR is not significantly different from Desktop AR on these two factors, it was

considered by subjects to be more intuitive, more comfortable to use and generally preferred. Furthermore, one should keep in mind that in the OR, the IGNS monitor may be much further away or less conveniently positioned due to other equipment, which could deteriorate performance or Desktop AR. Thus, we believe that even though our study was limited to a lab setup, the findings we have made would translate to the OR, similarly to how Cabrilo *et al.* [11] found that their AR system made a positive difference in two thirds of the clinical cases and a major improvement in 17% of the cases.

Although we did not quantitatively measure the difference in accuracy of the tracings between conditions, we believe they should be comparable, as Tabrizi *et al.* [6] and Cabrilo *et al.* [11] have shown that AR is not less accurate than the traditional systems. There is however a need to do a thorough study of accuracy of craniotomy planning between conditions. This will be done in future work.

The two most frequent negative comments that we received concerning our system were that there was some lag in the video feed and that the small size of the screen was making it harder to be precise. Those two limitations will be lifted in a future version of our system. The first one, caused by bandwidth limitations and network latency, will be greatly diminished when using compressed images. The second one will be solved when porting our system to a newer tablet device such as an iPad.

This work was motivated by feedback from a surgeon who has previously used Desktop AR in neurosurgery and wanted to be able to walk around the patient and see the location of relevant anatomy below the surface of the skin and skull during craniotomy planning. On presenting the prototype system to the surgeon, he commented that as well having the tumour projected, it would be useful to include vessels, gyri and sulci to further facilitate planning a resection approach. Furthermore, he commented that being able to look at the AR view on the mobile phone could

assist in teaching and allow easy discussion with residents in terms of surgical plan and approach. Given this feedback, in future work, we plan to add more features to our AR system, so that the surgeon can interact with the view on the phone, for example by turning anatomy of interest on and off and going through slice views mapped in depth to the real image.

3.7 Conclusions

In this paper, we examined the effect of *in situ* AR, desktop AR and traditional navigation on attention shifts between different IGNS displays and the patient. The results show that tumour outlining with AR systems takes less time and and requires fewer attention shifts than with a traditional navigation system. It is not clear that Mobile AR performs better on these two factors than Desktop AR, but it is clear that users find it more intuitive and comfortable. Reducing the disconnect between the AR display and the scene of interest does have an influence on the ease-of-use of the AR navigation system.

In future work, in addition to porting to iPad and compressing images, we also intend to bring the system into the OR to test it in its intended environment and with its intended users. As attention shifts have been shown to impact accuracy, we will further study the effect of the system compared to traditional image-guidance on the accuracy of different surgical tasks.

CHAPTER 4

Gesture-based registration correction using a mobile augmented reality image-guided neurosurgery system

Preface

In Chapter 3 we studied how mobile devices could have a place in the technological ecosystem of neuronavigational guidance. In this chapter, we directly apply those findings and present a novel method to use gestures on a mobile device to intraoperatively correct patient to pre-operative imaging registration error. This method builds on the advantages in terms of interactivity and intuitiveness of mobile

devices found in Chapter 3. Specifically, the surgeon uses typical touch screen gestures (*e.g.,* pinch and pan) to align the augmented reality patient anatomy with the real anatomy of the patient on the mobile system. Thanks to mobile devices' portability and possibility to drape in a sterile bag, the method enables the surgeon to make corrections to registration themselves, alleviating the need for a technician to be present. This can potentially streamline and shorten procedures and thus improve outcomes. The method can also be used in conjunction with other more automated brain shift correction techniques to make them more robust. The method developed is then used in the complete system development and assessment of MARIN (Mobile Augmented Reality Interactive Neuronavigator) described in Chapter 5. This paper was presented at the Augmented Environments for Computer Assisted Interventions (AE-CAI) workshop and published in a special issue of Healthcare Technology Letters.

Abstract

In image-guided neurosurgery, registration between the patient and their pre-operative images and the tracking of surgical tools enables GPS-like guidance to the surgeon. However, factors such as brain shift, image distortion, and registration error cause the patient-to-image alignment accuracy to degrade throughout the surgical procedure no longer providing accurate guidance. We present a gesture-based method for manual registration correction to extend the usage of augmented reality (AR) neuronavigation systems. Our method, which makes use of the touchscreen capabilities of a tablet on which the AR navigation view is presented, enables surgeons to compensate for the effects of brain shift, misregistration, or tracking errors. We tested our system in a laboratory user study with 10 subjects and found that they were able to achieve a median registration RMS error of 3.51 mm on landmarks around the craniotomy of interest. This is comparable to the level of accuracy attainable with previously proposed methods and currently available commercial systems, while being simpler and quicker to use. The method could enable surgeons to quickly and easily compensate for most of the observed shift. Further advantages of our method include its ease of use, its small impact on the surgical workflow and its small time requirement.

4.1 Introduction

In neurosurgery, surgeons treat different disorders which affect the brain, spinal cord, peripheral nerves, or cerebrovascular system. In order to do so they first diagnose the disease and make surgical plans using preoperative images, such as magnetic resonance images (MRI) or computed tomography (CT). Having access to the spatial location and extent of a lesion (*e.g.,* a tumour, arteriovenous malformation, etc.) is crucial to the success of a surgical procedure. Providing surgeons with this type

of information in the operating room has been one of the driving forces behind the development of image-guided surgery (IGS) systems. These systems have enabled more precise and minimally invasive surgeries compared to conventional surgical techniques [85].

In image-guided neurosurgery (IGNS), accurate and fast optical tracking systems, and registration of preoperative images to the patient allows for the real-time mapping and visualization of surgical tool positions and orientations with respect to preoperative images, thus guiding the surgeon; similar to a GPS system guiding a driver. One shortcoming of this type of traditional neuronavigational guidance is that the surgeon must shift their attention away from the patient and the surgical field to look at the guidance images on a computer display positioned outside the sterile area. Such shifts can disrupt surgical workflows and be detrimental to the task at hand (see Chapter 3). Recent work in augmented reality (AR) [11, 61, 6, 128] has addressed this issue by providing efficient and intuitive visualization of the complex 3D patient anatomy within the context of the live view of the operative field (see Figure 18).

The second shortcoming that remains with traditional IGNS systems is the continued loss of patient-to-image registration accuracy which degrades throughout the surgical procedure. Commercial IGNS systems show initial landmark registration accuracies between 2.7 mm and 6.2 mm, with a median of 4.0 mm, according to a meta-analysis done by Stieglitz *et al.* [116]. While this is true immediately after the registration procedure, this level of accuracy is no longer observed as soon as the craniotomy is performed. As summarized by Gerard *et al.* [39] in a recent review on the problem of brain shift, most of the studies measuring the problem reported mean shifts in the range of 1-10 mm and maximum shifts in the 10-30 mm range, with up to 50 mm [91] of shift. This makes brain shift the largest contributor to registration error; much higher than that of all other sources of errors combined,

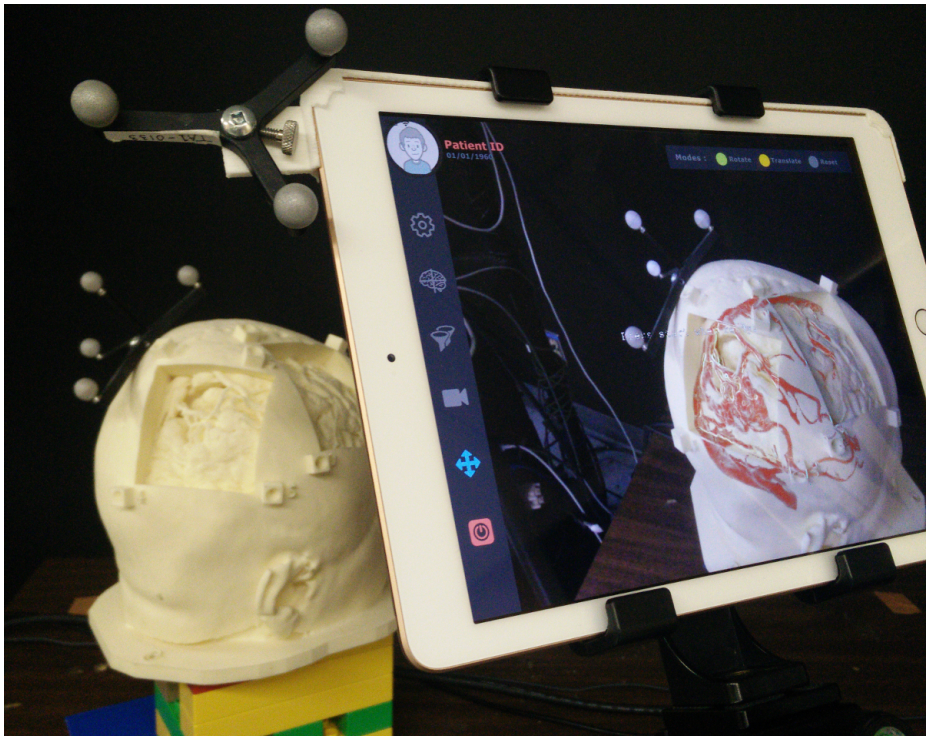

Figure 18: System set-up: the 3D printed phantom, trackers and the on screen live view of the phantom with vessels extracted from preoperative CTA are shown.

including errors arising from technical inaccuracies in the tracking, distortion in the preoperative images or initial registration error. Its impact is so large that many surgeons use IGNS systems to approach a surgical target but stop using it during the procedure, when the registration accuracy has degraded too much.

Albeit being the biggest limitation affecting IGNS systems, there is still no truly satisfying solution to solve the problem of misregistration. Brain shift is a complex phenomenon with multiple causes, making it hard to compensate for in an automated fashion. Many attempts have been proposed using either intraoperative imaging [93, 110, 105, 96], where intraoperative images are re-registered to preoperative ones, or biomechanical models [88, 83], where the aim is to predict the expected displacement using a patient-specific physical model. While these methods show promise, it is also clear that the road ahead to make these methods more robust, more general and less

sensitive to the occasional sparseness of intraoperatively acquired images and data will be long and strewn with pitfalls.

While there is much research on using intraoperative imaging or modelling to account for brain shift, simpler methods that aim to give the surgeon control over the registration not only at the beginning of surgery, but also during surgery, have not fully been explored. In this paper, we present a method to rigidly re-register images at any point in surgery using touchscreen gestures (*i.e.,* panning, rotation) on a tablet showing an augmented reality view of the surgical scene. The system allows the user to both translate and rotate the virtual preoperative patient images (visualized using AR) to the actual real-time images of the surgical scene. Rotation of the images is done with two fingers around the optical axis of the tablet's camera and translation is done with one or two fingers parallel to the camera image plane. Since the mobile device can very easily be moved around the patient, the user can translate in any plane and rotate around any axis. It thus gives access to the full range of rigid transformations to the surgeon. This method does not aim at replacing more complex non-rigid registration correction methods, such as FEM modeling or intraoperative imaging, but rather at complementing them. Our method is much simpler to use and has a negligible footprint intraoperatively, both time-wise and resource-wise. Thus, it could be used to make a quick rigid registration correction when time or resources are limited.

4.2 Related Work

The method we propose has a similar goal to the ones recently presented by Drouin *et al.* [24] and Kantelhardt *et al.* [57], which to our knowledge are the only manual registration correction methods in the literature. Drouin *et al.* proposed a method to perform manual re-registration by using a tracked pointer to trace vessels using

both the virtual patient data in an AR view (where the live video was captured by a neurosurgical microscope and the AR visualization was displayed on the computer monitor of the IGNS system) and on the actual patient cortex. Given the two sets of user defined vessel line traces, ICP (iterative closest point) was used to find a rigid transform between the points on the lines in order to correct for registration errors. Their results showed that users were able to correct for most of the registration error for medium to high shift cases, but degraded the registration for smaller shifts. They obtained a mean RMS error after correction of 4.06 ± 0.91 mm.

In Kantelhard *et al.*'s work, the user can translate the pre-operative patient images in x and y directions using arrows on the computer screen. Kantelhardt's method allows only for translation in the plane of the microscope, which is a strong limitation of the method. Considering this last point, we will use mostly Drouin *et al.*'s method as a basis for comparison to our method.

Our proposed method, contrary to Drouin *et al.*'s, does not make use of the tracked surgical pointer and offers *in situ* AR visualization. While it is true that for microscope-guided navigation the surgeon might already have the tracked pointer handy, it would not necessarily be the case for mobile AR-guided navigation. Indeed, with an *in situ* AR display, pointer free navigation is possible and has been shown to positively affect the surgical workflow by avoiding disruptions and limiting the number of times the surgeon must use the pointer to correlate guidance images to the patient [107]. We believe, therefore, that in this context our method would integrate more seamlessly into the surgical workflow and be more intuitive and simpler to use.

In this study, we tried to assess if our method would enable a similar accuracy as Drouin *et al.*'s, while using a simpler interaction method. Additionally, advantages over previous methods highlighted by Drouin *et al.* still hold for our method, namely: results are robust since they make use of the surgeon's extensive knowledge of the

anatomy; the method is quick and easy to use, making it possible to re-register as often as needed during a procedure; and the transform found with our method could be used as a starting point for more complex and automated methods, thus reducing the risk of them getting stuck in local minima. Further advantages of our method over previously proposed ones are: (1) the ability to easily compensate for shifts in all directions, not only in-plane ones (since the device can very easily be moved around and has much more freedom than the microscope); and (2) the use of simpler gesture-based interactions directly on the mobile device where the AR is displayed, which allows the surgeon to do the registration without relying on a technician to interact with the navigation system, nor using the surgical pointer.

4.3 System Description

Our mobile IGS system is comprised of a Polaris Tracking System (Northern Digital Technologies, Waterloo, Canada), an iPad (Apple Inc., Cupertino, USA) with MARIN: Mobile Augmented Reality Interactive Neuronavigator, a developed AR App, the IBIS (Intraoperative Brain Imaging System) Neuronav open-source platform for image-guided neurosurgery [25], along with a wireless router to relay data to and from the iPad to IBIS. The router through which both devices send video frames and commands is a TP-Link TL-WR810N, which uses the IEEE 802.11n wifi standard. The mobile application was built upon improving the system presented in Chapter 3. IBIS runs on a desktop computer with a i7-3820 3.6 GHz CPU, NVIDIA GTX670 GPU, ASUS PCE-AC55BT (intel 7260 chipset) wireless PCI card and Ubuntu 16.04.4 LTS (with the latest available wireless drivers (iwlwifi 4.13.0-43). The iPad used is an A1893 (6th generation) model, with the Apple A10 Fusion system-on-chip 64-bit architecture, an 8.0 MP camera and iOS 11.3. The iPad was outfitted with a passive tracker that was attached to a custom 3D printed case (see Figure 18).

The IBIS Neuronav package comes with plug-ins for tracking, patient-to-image registration, camera calibration, and the capability to use augmented reality visualization by capturing a live video stream from a microscope or video camera and merging it with preoperative images on the monitor of the system itself. In our work, we extended the IBIS Neuronav system to allow for augmenting an image, not only on the monitor of the system but on a mobile device (*i.e.,* a smartphone or tablet) that captures the surgical field of view, thus allowing for *in situ* AR visualization, and the integration of gestures. The system set-up as used in the lab on a 3D printed phantom head is shown in Figure 19.

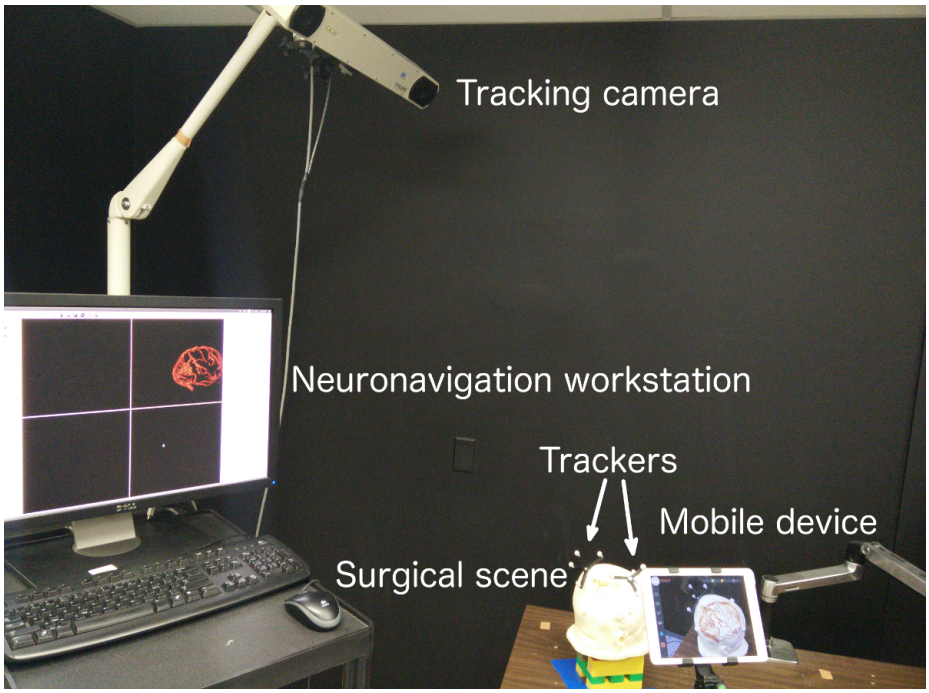

Figure 19: System setup: tracking camera in the top-left corner, IBIS workstation in the lower-left corner and tablet in the lower-right corner with phantom.

To make use of IBIS' existing functionality, the mobile device (*i.e.,* iPad) serves merely as a camera and display. The costly computations are handled by the desktop on which IBIS runs. In order to create the AR view, we first calibrate the camera of the tablet. Calibration (intrinsic and extrinsic) is done using a modification of Zhang's

camera calibration method [137], followed by a second optimization procedure to find the transform from the tracker to optical center of the camera (for more details the reader is referred to [25]). The average camera calibration reprojection error obtained was 0.77 mm. Patient-to-image registration is done using landmark registration, for which we obtained a RMS error of 0.69 mm.

IBIS receives the positions of the iPad and patient from the tracking camera, then computes the relevant transformations and renders the virtual objects from the camera's view point. The rendered virtual objects are then sent using OpenIGTLink [120] over the local area network to the tablet, and blended with the live video feed using OpenGL ES 3.0 and GLSL. The Qt framework (version 5.10) was used in the design of the AR application user-interface, and also used in the communication with the mobile camera.

A screenshot of the interface of the MARIN app is shown in Figure 20. Options for the registration correction module are displayed in panel at the top-right corner. These options serve as switches to turn the rotation and translation mode on or off and to perform either both at the same time or only one at the time if needed. The last button in this panel is a "reset" button, which allows the user to restart the correction procedure from the initial misregistered image.

The touchscreen gestures registered by our app are (1) panning and (2) pinching with rotation, both of these are common gestures used in mobile interfaces and should thus be intuitive to users. Panning results in a translation of the objects in the plane of the camera image. Pinching with rotation, also sometimes simply referred to as rotation, results in a rotation of the object around the optical axis of the camera. Both gestures result in a transform that is directly applied to the virtual object. Users therefore, have access to the full range of rigid transforms by compounding translations and rotations from different perspectives.

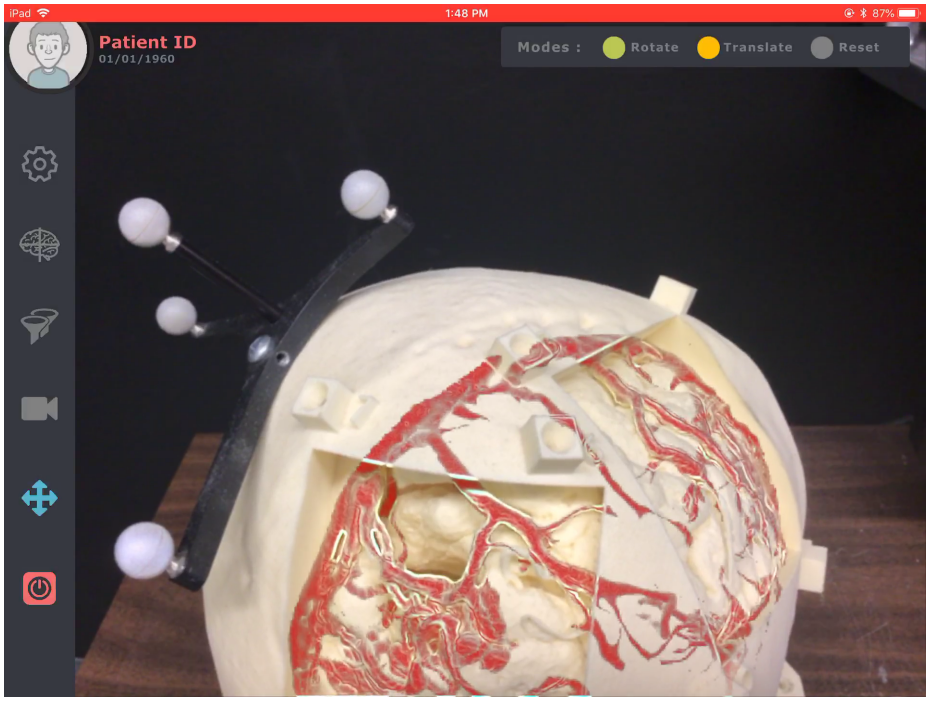

Figure 20: Screenshot of the iPad AR application during use. Options for the registration correction module are displayed in the top-right corner. The green and yellow switches allow the user to turn rotation and translation modes on or off to perform either both at the same time or only one at the time. The last button on the right is a reset button that allows the user to restart the correction procedure from the initial misregistered image.

4.4 Methodology

Testing of the usability and accuracy of our system was done in a user study where participants were presented with a misregistered AR view and were asked to correct the registration with our system. More information on the study design and sample is given in section 4.5.

The phantom on which the study was performed is a 3D nylon printed head whose model was extracted from MRI and CT angiography (CTA) data of a patient operated on for an arteriovenous malformation (AVM). Two simulated craniotomies expose the cortex and surface vessels: one in the posterior parietal lobe, centered on the AVM, where multiple vessels are visible and the other one in the temporal

lobe, where only few vessels were visible. At the corners of each of the square-shaped craniotomies are four landmarks, which were used for the initial registration, and the determination of the RMS error. See Figure 18 for an example view of the system during the experiment.

A close-up of vessels near the AVM in the first craniotomy is shown in Figure 21. In the bottom two images, an example of registration correction performed with the system is given. On the left, we can see vessels translated from their correct position, while on the right they have been realigned. Note that edges in the camera image are detected and used to modulate the blending coefficient between the real and virtual images in the AR view. The augmentation opacity on edge pixels is reduced to allow the user to see boundaries of real objects. These boundaries, as we can see in the bottom-right image of Figure 21, should line up with the boundaries of the virtual objects when images are properly registered. This cue could thus be used by users to guide them during the manual registration correction procedure.

Note that in the user study we performed, the displayed mesh was that of only surface vessels of the relevant hemisphere. It was thus similar to what would typically be shown to a surgeon for an AVM procedure on an AR overlay.

The measured outcomes of the study were registration RMS error after re-registration and percentage correction from the initial misregistration. Further, total time to re-register the images and the number of times subjects moved the tablet around the phantom to look at the misregistraton from different perspectives was measured.

4.5 Experiment

A protocol as similar as possible to Drouin *et al.*'s was chosen, in order to be able to compare our method with their pointer based re-registration method. As in Drouin

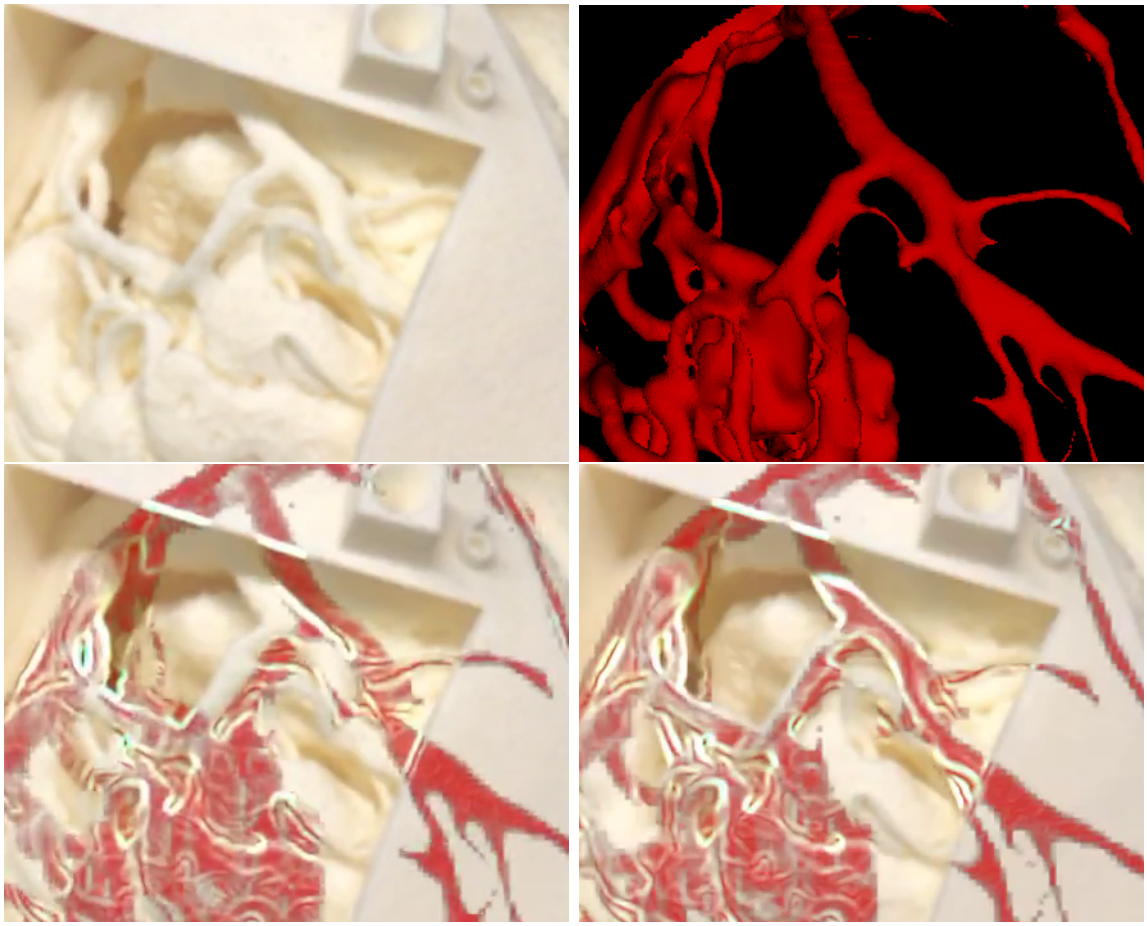

Figure 21: Augmented reality visualization. Top-left: Phantom without augmentation. Top-right: Virtual image of vessels. Bottom-left: Augmented reality view where the volume is misregistered, as seen on the tablet. Bottom-right: Augmented reality view where the volume has been re-registered, as seen on the tablet.

et al.'s validation, subjects were presented with 5 misregistrations to correct, all of the same preoperative images and on the same phantom. Initial registration offsets were the same as in Drouin *et al.*'s evaluation and are given here in Table 1. Offsets were generated by translating in plane with a given amplitude, in a random direction and rotating around the optical axis with a given angle, in a random direction. Both translation amplitude and rotation angle decreased with every trial.

Our study sample was composed of ten subjects (aged 24-57 (median 29), 7

Trial no.	Translation amplitude (mm)	Rotation amplitude (deg.)
1	15	5
2	10	4
3	6	3
4	3	2
5	1	1

Table 1: Initial misregistration offsets in rotation and translation from correct position per trial.

males and 3 females). They were students, professors, engineers and neurologists, all working in either biomedical engineering, computer science, or medical image analysis. All subjects were already familiar with neuronavigation systems and augmented reality. Nevertheless, they were re-familiarized with the concepts and the method's intended usage prior to the study. Subjects were then briefed on the system functionality, interaction modes and study procedure. After the instructions, but before the beginning of the trials, subjects were presented with a correctly registered AR view and given time to become familiarized with the system. They were asked to use the touchsreen gestures to move the virtual images around and to re-register them as accurately as possible. This pre-test trial served as a practice and to help reduce potential learning bias. It was emphasized to the subjects that the goal was to be as precise as possible and make the best registration correction as possible, for which they could take as much time as they felt was necessary.

After the evaluation, subjects were asked to fill out a questionnaire consisting of demographic questions, the System Usability Scale (SUS) questionnaire [10] and additional questions on the method's usability (a copy of which can be found in Appendix A). The SUS is a standardized usability questionnaire consisting of ten five points scale questions. It has been used on a large number of applications and interfaces, for which data is available. It is therefore a well established test to estimate a system's usability and compare against other systems. A system scoring above 68

is considered usable and above average.

4.6 Results

Before reporting any measure, it is worth mentioning that two individual trials were flagged as outliers. Their registration RMS error after correction were respectively 7.7 and 10.5 standard deviations away from the distribution mean. Considering that these two trials (coming from two different subjects) were much further away from normal performance than any other trial both across all subjects and within each of the subjects, we believe that there was an experimental error or technical issue on those trials. We therefore removed these two trials from all further analysis.

The median registration RMS error after correction for all subjects on all trials was 4.13 mm. Interestingly, if looking separately at the landmarks around the two distinct craniotomies, we observe that the median registration RMS error distances are significantly different: 3.51 mm (for the craniotomy of interest where the AVM and many other vessels were visible), and 4.88 mm for the second cranitotomy. A one-way ANOVA comparing the registration RMS error of the landmarks around each of the craniotomies shows that there is a statistical difference between the RMS error means ($p = 0.0006$). Landmarks around the prevalent craniotomy will be used in further analysis.

Although we believe that the median is a better outcome measure in this case since the distribution is not Gaussian, we still computed the mean, in order to compare against Drouin *et al.*'s method. We obtained a mean registration RMS error after correction of 3.84 mm with a standard deviation of 1.34 mm, comparable to Drouin *et al.*'s error of 4.06 mm with SD of 0.91 mm. The RMS error after correction against the initial misregistration RMS distance, for comparison with Drouin *et al.*'s method, is

shown in Figure 22. Distribution of the points and linear regression fit are comparable to what is reported in [24].

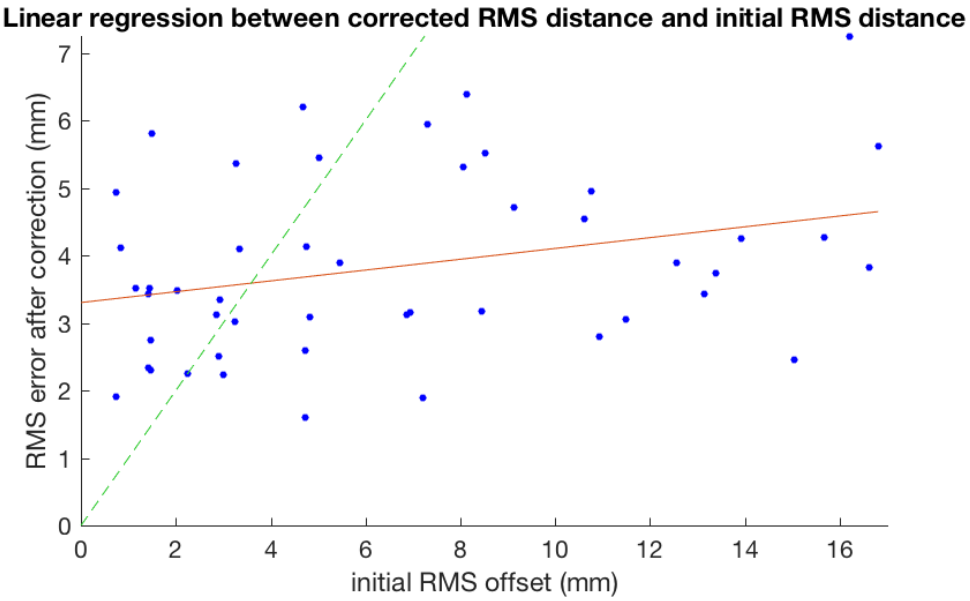

Figure 22: Data points for all trials and all subjects presenting relationship between the initial registration RMS error and the registration RMS error after correction (blue dots) with its linear regression fit (red line) and the zero correction partition (green dashed line).

A more telling measure of performance is the percentage of misregistration RMS error that subjects were able to correct for as shown in Figure 23. This graph makes it easier to distinguish between negative and positive performance. This difference can also be seen in the previous graph, where points either lie under or below the zero correction partition (green dashed) line. As we can see from the graph, when the initial misregistration was under the median RMS error of 3.51 mm subjects had a negative performance, degrading the registration error. However, in looking at only medium to high shifts (*i.e.,* an initial misregistration RMS greater than 4 mm), which would be the target usage for the method, subjects achieved a median percentage correction of 55%.

The time taken to complete the correction ranged from 11 to 161 seconds,

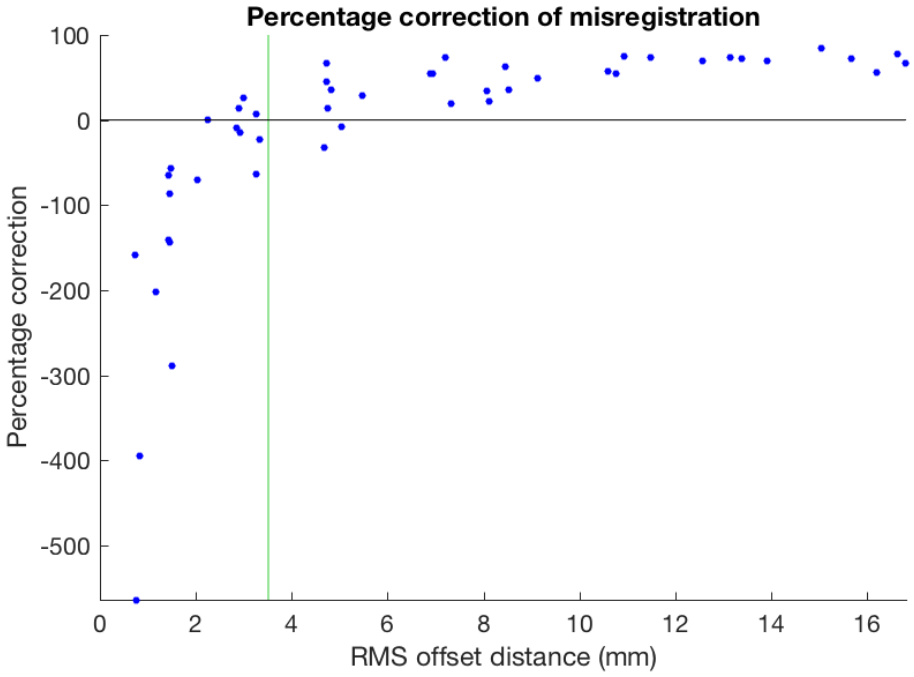

Figure 23: Data points for all trials and all subjects presenting relationship between the initial registration RMS error and the percentage of RMS error corrected (blue dots). The median RMS error after correction is indicated by the green line. Data points under the grey line (equivalent to 0% correction) indicate negative performance, *i.e.,* the subject made the registration worse. Data points above the zero correction partition line indicate the percentage by which the subject improved the correction.

with a median 52 seconds. The number of tablet displacements varied a lot from subject to subject (range 0-10) and seemed to be more of a personal preference. There were no significant interaction effects found between time taken, number of tablet displacements, and accuracy. The Pearson's correlation coefficients for time and number of displacements with final registration RMS error are 0.13 and 0.05 respectively, indicating that neither have influence on the accuracy. Looking at those two factors against initial registration error, the Pearson's correlation coefficients are 0.02 and -0.15 respectively, showing here that the initial offset also does not influence the required time to complete the correction or the number of required movements.

In terms of qualitative evaluation, a few comments came out of the usability survey

we gave our study participants. Three shortcomings in the interaction were perceived by users as having a negative impact on the attainable accuracy: jitter, latency and sensitivity. Jitter is an artifact of the tracking system that causes the AR overlay to shift slightly due to inaccuracies in the infrared reading. It is entirely dependent on the tracking system and is thus not particular to our method. Latency is the time between the moment the subject makes a gesture and the moment the images get updated. This latency was caused by the compounded effect of many factors, including network transfer of the command and of the augmentation images and was in the hundreds milliseconds range. Sensitivity is considered the ratio between gesture amplitude and resulting transformation amplitude. Despite these shortcomings, the mean SUS score across subjects for our system was 70.5, which is above the average of 68 indicating a usable interface and system.

4.7 Discussion

The first thing that stands out when looking at Figure 23 is the clear distinction of subjects' performance between trials where the initial RMS registration error is above and below 4 mm. When above, subjects were almost systematically able to correct for the most part of the shift. Although, when below, subjects usually deteriorated the registration accuracy. This would seem to indicate a bound on the attainable accuracy with our system in its current state. Looking back at Drouin *et al.*'s results, a similar trend can be noticed. We posit therefore that this is perhaps a limitation of the study design. Further, it is not clear that this deterioration would translate in clinical practice. In the setting of this and Drouin *et al.*'s study, subjects had to try to correct the registration, even when it was only very slightly offset. In clinical practice, we believe that surgeons would only use the correction method when they feel the preoperative images have suffered a visually significant shift, typically several

millimeters, and the registration is no longer accurate enough for guidance. Thus, in those cases where the initial RMS offset was below 2-4 mm (*i.e.,* a typical commercial system registration error), for most surgeons probably, it would have been deemed usable and re-registration would not be done.

The second distinction to raise between our study and clinical practice is the user's prior knowledge. We only enrolled participants who were familiar with AR, IGNS, and medical imaging data, which made them similar to surgeons in that respect. However, many of them were not necessarily experts of neuroanatomy, contrary to surgeons. Furthermore, they had not done preoperative planning on the images and thus were not familiar with them, unlike in real clinical cases where the surgeon would already be familiar with the anatomy. Considering those two points, we believe that the median accuracy obtained from our subjects in this study could be considered to be a lower bound on what surgeons would be able to achieve.

Our method produced a mean registration RMS error of 3.84 mm compared to Drouin *et al.*'s result of 4.06 mm. While our result is slightly better, it is hard to say if our method really offers better accuracy considering the relatively small sample size of both studies (10 in our case and 5 in the case of Drouin *et al.*). A direct comparison of both methods with a larger sample size would have to be done in order to gain a better understanding of the potential differences in attainable accuracy. With the current data, we can at least say that they have comparable levels of accuracy. Additionally, the linear regression fit for both samples seems to also be comparable.

We posit that the larger error on landmarks further away from the craniotomy of interest resulted from the relatively large difference that a slight error on the rotation causes on points far from the center of rotation. This error, however, may not be clinically as important as ones close to the site of interest. This is in line with what real use cases would be. Surgeons would typically be interested in targets

close to the site of the craniotomy and would not necessarily be much affected if a slight rotation caused landmarks on the opposite side of the cranium to be slightly misplaced. Further, it is deemed that in real cases surgeons would only use the system if they felt that enough features were visible to allow for the correction to be done. Additionally, surgeons who have better knowledge of the anatomy may be able to pick up smaller features to match than non-experts.

Another interesting point to come out of the data is the very low time taken to perform the correction. All corrections, even though subjects were instructed to take as much time as needed to perform the best registration possible, were done under three minutes and the median time to correction was under a minute. The short time requirement as well as the fact that our method can be performed by surgeons themselves without intervention of a technician hints that the disruption caused by our method to the procedure is minimal. Additionally, since the tablet can be draped in a sterile bag, our method can easily be used throughout surgery and as often as needed. Although, and as alluded to in the introduction, this method would not substitute itself to more complex non-rigid brain shift compensation methods, but it could be used early on and in between uses of those methods. Indeed, non-rigid compensation methods are computationally expensive and require usage of a form of intraoperative imaging, whether it be ultrasound, intraoperative MRI or CT, which makes them much more heavy to use both time-wise and resource-wise. Thus, the real potential of our method lies not in contrast with more complex methods, but in complementarity with them.

As discussed above, time was not correlated with registration accuracy or amplitude of initial shift. This would hint at two interesting findings. First, it provides an incentive to use this method to compensate for large shifts since the time requirement stays the same and is very small regardless of the amplitude of the shift.

It further motivates the use of the method as a first correction before using a more complex and automated method to compensate for most of the rigid misregistration in less than a minute in order to ensure that an automated method would not get stuck in a local minima. Second, regarding the perhaps surprising absence of correlation between time spent and final accuracy achieved, we posit that there is an upper bound in terms of registration accuracy achievable with the method. However, this bound can be achieved within a short amount of time.

Regarding subjects' comments, the two factors specific to our method that were seen as negatively impacting accuracy, namely latency and sensitivity, could be improved in a future version of the system. Network bandwidth is currently the largest contributor to latency. Latency could thus be reduced by optimizing compression in image transfer and upgrading the router and network card of the workstation to the newer 802.11ac wifi standard. Sensitivity would also be increased, but perhaps even better would be giving the users control by allowing them to set the sensitivity themselves from within the application's interface. This would allow users to quickly move the volume in its approximate position and then increase sensitivity towards the end to gain finer control and potentially reach a better registration accuracy.

Another interesting comment from users was about the axis around which the volume is rotated. In the present version of our method, the volume is rotated around the optical axis of the camera. Some users however, mentioned that it may be easier if the rotation was done around an axis parallel to the optical axis of the camera, but translated in the plane of the camera to pass through the center of the volume. We did not picture this as a potential problem since we thought users would usually have the object of interest more or less in the center of the screen, in which case those two axes would coincide. However, this was not always the case.

A final comment raised by participants which seemed of particular interest for

future revisions of our proposed method is the ability to translate in the z direction without having to move the tablet. In the current version, if a user wants to translate in z, they must first physically move 90 degrees around the patient's head, and translate in plane by the desired distance (in what is now the x or y direction), then return to the initial position to verify that the transformation is correct. This can be time consuming and often requires moving the tablet many times in order to obtain the desired correction. It was suggested by users to add a pinch gesture to translate along the z direction of the current view, thus enabling quicker correction with less movements.

The usability score of 70.5 we obtained for the SUS places our system slightly above the average for the test of 68. While not very high, it certainly shows, that even in the present state, the system is usable. Although, it can be assumed that after integrating user's comments and suggestions, a further version would be much more usable and also potentially perform better in terms of accuracy.

4.8 Conclusion

This study shows that using gestures on a mobile device to correct registration error of an AR IGNS is a valuable option. Users were able to achieve accuracy comparable and even slightly better than previously proposed methods. The collected data further suggests that a sufficient number of image features (vessels, in our study) would be necessary in order for subjects to achieve a valid correction. For other types of procedures where angiography data is not available, cortical features could be matched. Results also show that, similarly to previous methods, the relevance of the method varies with the amplitude of the initial misregistration. Indeed, for medium to large initial shifts (*i.e.,* more than 4 mm), subjects were able to account for most of the error, but for smaller shifts, subjects were usually unable to improve the registration.

The present method's usability could be improved by integrating comments from users such as adding interaction modes and tweaking interaction parameters. It could be hypothesized that enhancement in usability would lead to enhancement in achievable precision. Additionally, more image data sources could be added to the AR view and combined in a coherent model, such as the cortex surface in addition to vessels. Considering how well acquainted with the anatomy surgeons are, the more data that is made available to them the more likely they would be to find helpful features for the correction, conditional to maintaining a clear and uncluttered view of the data. Although the iPad AR system itself has been brought into the operating room for initial testing where we have received positive feedback, we have yet to test the registration feature. Once the improvements mentioned above are implemented in a future revision of the method, they will be tested with surgeons during real clinical cases in the operating room.

CHAPTER 5

MARIN: an open-source mobile augmented reality interactive neuronavigation system

Preface

After assessing mobile device's potential for neurosurgical guidance use in Chapter 3 and devising and testing a gesture-based registration correction method in Chapter 4, in this chapter we present a complete mobile neuronavigation system implementing the AR window paradigm, displayed on a mobile device. This system makes use of knowledge gained in Chapter 3 and integrates the method devised in Chapter 4. This complete system implements an *in situ* augmented reality visualization, as well as replicates the functionality of a commercial system, all displayed on a mobile

device. The user can switch between the different visualizations and customize the displayed information intraoperatively, without requiring assistance from a technician. This system improves relative to the previous state-of-the-art in all respects. It is more versatile and offers much higher performance than previously proposed system, making the system clinically relevant. As well, to increase accessibility, the code is released under an open source license. Furthermore, this chapter's work, not only presents a novel system, but as well pushes further the knowledge from Chapter 3. In this chapter, a thorough assessment of the devised system used in its different modes of operation, enabled us to again compare AR vs non-AR visualizations, but this time with respect to subject accuracy for target localization. This paper was presented at the Information Processing in Computer-Assisted Interventions (IPCAI) conference and published in the International Journal of Computer Assisted Radiology and Surgery. Finally, an extension to the initially published system is presented at the end of the chapter as an addendum. This extension adds an intraoperative AR display of hemodynamic guidance feature to the system. This extension was presented at the Computer Assisted Radiology and Surgery conference (CARS) and published as an abstract in the International Journal of Computer Assisted Radiology and Surgery.

Abstract

Purpose: Neuronavigation systems making use of augmented reality (AR) have been the focus of much research in the last couple of decades. In recent years, there has been considerable interest in using mobile devices for AR in the operating room (OR). We propose a complete system that performs real-time AR video augmentation on a mobile device in the context of image-guided neurosurgery.

Methods: MARIN (Mobile Augmented Reality Interactive Neuronavigation) improves upon the state-of-the-art in terms of performance, allowing real-time augmentation, and interactivity by allowing users to interact with the displayed data. The system was tested in a user study with 17 subjects for qualitative and quantitative evaluation in the context of target localization and brought into the OR for preliminary feasibility tests, where qualitative feedback from surgeons was obtained.

Results: The results of the user study showed that MARIN performs significantly better both in terms of time ($p < 0.0004$) and accuracy ($p < 0.04$) for the task of target localization in comparison to a traditional image-guided neurosurgery (IGNS) navigation system. Further, MARIN AR visualization was found to be more intuitive and allowed users to estimate target depth more easily.

Conclusion: MARIN improves upon previously proposed mobile AR neuronavigation systems with its real-time performance, higher accuracy, full integration in the normal workflow and greater interactivity and customizability of the displayed information. The improvement in efficiency and usability over previous systems will facilitate bringing AR into the OR.

5.1 Introduction

Neuronavigation has historically been at the forefront of surgical technological development due to the unique constraints of neurosurgery, which on the one hand

require maximum resection of the disease and on the other hand requires minimum damage to healthy or eloquent tissue [34]. Owing to this, accuracy is of paramount importance, and therefore neurosurgery has often inspired new innovations that attempt to refine guidance accuracy, allow for distinguishing between healthy and diseased tissue, and improve the intuitiveness of guidance visualization, in order to improve surgical outcomes. A historical review of neuronavigation was presented by Galloway and Peters [35].

One solution proposed to facilitate neurosurgical guidance has been augmented reality (AR). The first use of AR for neurosurgical guidance dates back to the late 80s [106]. More recently, with the advent of smart mobile devices, new means of displaying intraoperative AR views, through the AR window paradigm have been explored (*e.g.,* [128, 19]) for instance. The use of mobile devices for AR display in the clinical context has many advantages over other *in situ* AR setups. First, these devices offer an easy means to not only display information but also interact with it using the touchscreen. Thus, there is much more potential for interactivity than with projector-based AR systems for instance. Second, these devices are small and wireless, which make them easy to move around and explore the anatomy from different perspectives, allowing for easier planning and teaching [61]. Third, they are much more versatile than head-mounted displays (HMDs), which can be bulky and disruptive. Indeed, current HMDs have been shown to be poorly suited for surgical tasks [14]. Additionally, contrary to HMDs, mobile devices can be draped in sterile bags, allowing for continuous use throughout a procedure. Mobile devices can be stowed away and brought back in a matter of seconds or clamped to the bed for hands-free continuous guidance. In this paper, we propose the MARIN (Mobile Augmented Reality Interactive Neuronavigation) system, which enables more intuitive and interactive visualization in image-guided neurosurgery (IGNS). The

utility, usability, and performance of the developed system was evaluated in the lab and with feedback from surgeons.

5.2 Previous Work

A number of mobile AR systems have been proposed, implemented and/or tested in a clinical setting. Hou *et al.* [51] and Eftekhar *et al.* [29] both proposed phone apps that display the video feed of the camera overlaid with a previously selected slice of a preoperative CT or MRI. The system is set-up in the OR and the slice manually aligned with the patient intraoperatively, *i.e.,* the registration is done manually thus the accuracy relies on the operator. Owing to this the AR views are limited to be along a scan axis (*i.e.,* saggital, coronal or transverse), defined during preoperative planning. The systems from Deng *et al.* [19] and Watanabe *et al.* [128] consist of a tablet app showing presegmented structures overlaid on the live video feed. Contrary to previously mentioned systems, they make use of an external tracking system. This enables them to show the AR view from any angle around the patient, thus offering a significant improvement compared to previous work. At the same time these systems do not run in real-time, due to a large latency in image transfer. Watanabe *et al.*'s latency between tablet movement and updated AR information is about 400 ms. Deng *et al.* do not report latency, yet mention an alignment error caused by a certain degree of delay. This kind of latency, as shown by Sielhorst *et al.* [113], has a strong negative impact on task success. Even small discrepancies in time cause the impression of a larger discrepancy in space. Additionally, none of the current mobile AR systems take advantage of the touch screen of these devices to allow the user to interact with the displayed information.

In our previous work (Chapter 3), we developed a mobile AR IGNS prototype that worked on an Android smartphone. Similar to the above mentioned works, this

first prototype did not run in real-time and did not allow the user to interact with the system through the mobile device. This paper aims to address the limitations of our and previous systems. Namely, the main improvements over the state-of-the art is MARIN's real-time performance (50 ms average latency), full integration into the surgical workflow, greater customizability and interactivity. In an effort to accelerate development in the field and enable faster improvements, with the publication of this paper we are making the software available under an open-source license. We strongly believe that development of better healthcare technologies will happen through sharing and collaboration of teams from around the world.

5.3 System Description

In the following section, the implementation details of the developed MARIN system are described.

Figure 24: Photos of the system, as was used in test study described in section 5.4. On the left: the iPad running MARIN mounted on an adjustable tripod, the phantom and the tracking system mounted an a larger immovable tripod at the back, behind the table. On the right: the iPad running MARIN and the tracked pointer.

5.3.1 Hardware

MARIN, being an open source project and in an effort to make it as widely available as possible, is compatible with a broad range of hardware. The required hardware components are a mobile device, a desktop computer, a tracking system and a wireless router. The mobile device can be either a phone or tablet, running iOS or Android. The desktop computer may run Linux, OS X or Windows, so long as it has networking capabilities and a graphics card. The router should support at least the 802.11n wireless standard and the WPA or WPA2 security protocol. The choice to include a dedicated router in our setup was made in order to ensure low latency, while still respecting security concerns. The router offers a private network on which only devices from our system are connected ensuring security. Thus additional encryption, which would unnecessarily increase transfer latency, is not required. Further, this allows MARIN to not rely on any external network. However, if the system was to be used in an environment with access to a trusted network, this component could be omitted. Finally the tracking system can be any of those supported by PLUS [68], which includes systems with very diverse specifications, including some small and cheap ones, as well as bigger and more precise ones, making it applicable in a broad range of applications, environments and setups. These hardware choices ensure the widest availability and portability, allowing MARIN to be used in as many different settings as possible.

For our tests, the following hardware was used: a FusionTrack500 Tracking System (Atracsys LLC, Puidoux, Switzerland), an iPad model A1893 running iOS 12.4.1 (Apple Inc., Cupertino, USA), an Archer A10 router (TP-Link USA Corporation, Brea, USA), a desktop computer with a i7-6850K 3.6 GHz CPU, NVIDIA GTX 1080 GPU, Gigabyte GC-WB867D-I wireless PCI card, running Ubuntu 16.04 LTS. The iPad was outfitted with a passive tracker that is attached to a 3D printed attachment

bracket (Figure 24).

5.3.2 Software

MARIN makes use of the IBIS Neuronav open-source platform for IGNS [25], which comes with plug-ins to integrate tracking, do patient-to-image registration, camera calibration, and the capability to do augmented reality visualization (by capturing a live video stream from a microscope or video camera and merging it with preoperative images on the monitor of the system itself). In our work, we extended the IBIS Neuronav system to allow for augmenting an image on the screen of a mobile device that captures the surgical field of view. A custom built plugin (OpenIGTLinkSender) was contributed to the source code. The plugin enables images to be sent from the desktop computer running IBIS to MARIN, allowing for *in situ* AR visualization, through the AR window paradigm (not previously possible with IBIS). It should be noted that our system's usage is not limited to interaction to IBIS. Our mobile application could work in conjunction with other systems offering similar functionalities, for example SlicerIGT [121]. This is due to the fact that MARIN uses OpenIGTLink [120], an open protocol and library for message transfer between devices, specifically made for the context of IGS. It thus enables any system that supports OpenIGTLink to communicate with MARIN. Choices about software used in the system were made with the same goal in mind as for hardware support. We aimed at making the system as customizable and portable as possible. All libraries are open-source (PLUS [68], OpenIGTLink, OpenIGTLinkIO, libyuv and OpenH264), and can run on all common operating systems and hardware. The H264 codec used for image compression is also widely supported and has hardware acceleration on modern mobile devices, enabling shorter latency to be reached. MARIN's source code can be found at https://github.com/AppliedPerceptionLab/MARIN.

Figure 25: Block diagram representing interactions between system components and data flow.

To make use of IBIS' existing functionality, the costly computations are handled on the desktop computer. In order to create the AR view, the mobile device camera is first calibrated. Calibration (intrinsic and extrinsic) is done using a modification of Zhang's camera calibration method [137], followed by a second optimization procedure to find the transform from the tracker to optical center of the camera (details are given in [25]). This calibration step takes about 10 to 15 minutes to perform. In our surgical workflow, calibration is done in the laboratory prior to bringing the system into the OR. Thus it does not add time to the procedure itself and does not require involvement from either the surgeon or OR staff. It does however need to be performed before every operation. Tracking information is sent from the tracker to IBIS through a PLUS server. The relevant transformations are computed in IBIS and renderings of the virtual objects made, from the camera's view point. Using the newly developed plugin the rendered virtual objects are compressed with the OpenH264 library and sent using OpenIGTLink through a UDP socket over the local area network to the tablet on which the virtual data is blended with the live video feed using OpenGL ES 3.0 and GLSL. Modifications were made to the OpenIGTLinkIO library to support transport over UDP sockets. The use of the UDP transport protocol, as well as video compression allows for shorter transfer latency. A block diagram of system components, their interactions and data flow is shown in Figure 25. In addition to

image transfer from the desktop to the mobile device, other communication channels are opened between the two devices to send images from the device to the desktop and commands and status updates. These communication channels enable interactivity and more possible use cases. Since the camera view can also be sent from the mobile device to the desktop, it is possible to show the same AR view on the desktop, as well as on the mobile device, at the same time. This could be used for training and teaching as residents and other staff can see the same AR view as the surgeon, on the desktop screen. Channels for status and command transmission are also used to enable interaction with the AR view and see real-time information from tracking.

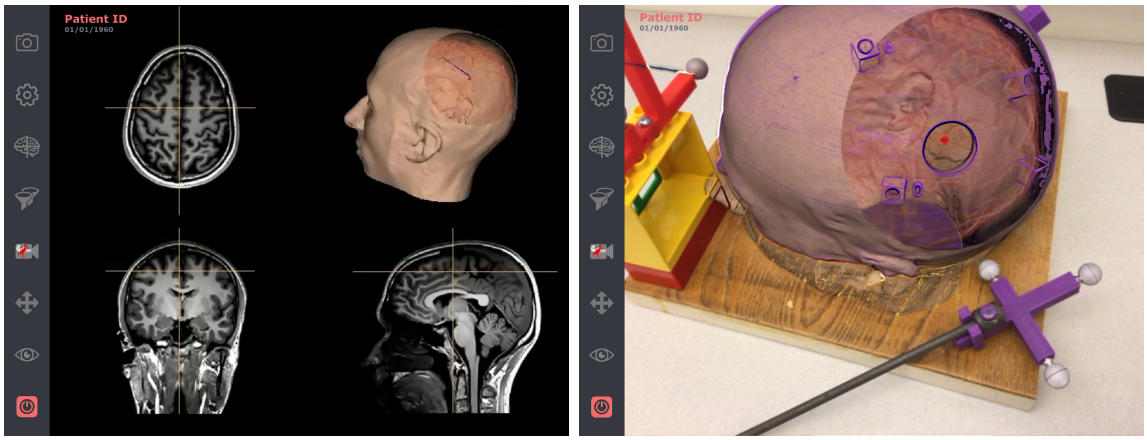

Figure 26: On the left: screenshot of the system when used in standard IGNS mode; On the right: screenshot of the system when used in augmented reality mode.

5.3.3 MARIN Interface

MARIN offers an intuitive interface, letting the user easily customize what they want to be displayed on the AR view, in real-time, during the procedure. The view can be switched between camera-only, virtual only or AR. Users can pick only an area in which they have an AR view, potentially reducing clutter in cases where there is a lot of data needed by the surgeon. Different structures from the AR view can be

switched on or off as desired, *e.g.,* skin, tumour, vessels, etc. Users can also choose from different navigation paradigms. They can switch between the standard IGNS view and AR view (see Figure 26) or show only the virtual view, showing only pre-segmented 3D models. MARIN also features a manual registration correction feature, which allows the user to realign the (virtual) patient pre-operative images with the real patient (see Chapter 4) using typical touchscreen gestures. All interactions with MARIN are done through simple gestures, making it intuitive and easy to operate, even if draped in a bag. Furthermore, MARIN can receive any type of image over the network through OpenIGTLink messages. This image could, for instance, be a static MR slice with highlighted structures, as was done by Hou *et al.* and Efthekar *et al.* MARIN can thus reproduce the behaviour of previously proposed systems, but also do much more. The interface was developed using the Qt framework (version 5.9.8) which allows the application to be easily compiled for different devices.

5.4 User Study

To evaluate MARIN's performance and assess its usability and accuracy, we conducted a phantom study in the context of target localization. This task replicates some common surgical tasks such as electrode and port placement or ventriculostomy catheter insertion.

5.4.1 Phantom Construction

A T1-weighted image of a healthy subject was segmented to extract the skin and cortical surfaces. The 3D model of the skin surface was used to represent a patient head and an 18 mm wide craniotomy on either side of the back of the head was added. Seven landmarks were added to the model for easy registration (inspired by

those used by Drouin *et al.* [23]). The model was printed using PLA filament on a Replicator 5th Gen (MakerBot Industries, USA). The printed phantom was affixed to a rigid board to which a marker was attached (Figure 27).

A gelatin brain phantom was inserted into the cavity of the 3D printed head. Gelatin was chosen for its ease of use and because it is the most widely used material for brain phantom construction [69]. We used a 10% weight to weight ratio, heating the solution to 40°C, casting it into a hemispherical mold and letting it set for 12 hours at 4°C. We chose a concentration closer to human muscles than normal brain tissue in order to increase stiffness and make it more difficult for subjects to move the pointer sideways once insertion had begun. This was desired, to simulate the conditions of a surgical setting more closely; surgeons are required to plan an insertion path and follow it once they start inserting. Moving laterally would risk causing unnecessary damage to healthy tissue next to the insertion path.

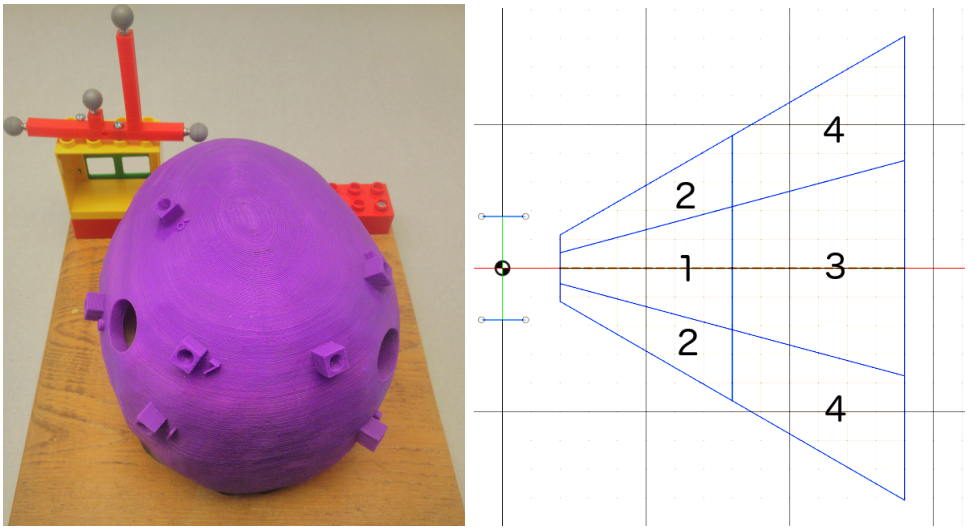

Figure 27: Experimental phantom design. On the left: the phantom seen from the back, with craniotomies on either side, fiducials and the marker behind. On the right: a depiction of the four difficulty zones used in the experiment. The craniotomy is located between the two blue lines on the left and is centered on the dot. Then the four zones are: 1: 0-15°from normal, 1-4 cm deep; 2:15-30°, 1-4 cm; 3: 0-15°, 4-7 cm; 4: 15-30°, 4-7 cm.

5.4.2 Task Description

Participants were asked to reach target points within the phantom with a tracked pointer as precisely as possible. The task was performed under two different modes: traditional IGNS view (displayed on the tablet) and the AR view from MARIN. The IGNS navigation system served as a baseline to assess MARIN's performance. Both systems were presented *in situ*, thus removing the potential confound of location of displayed information, since this has been previously shown to potentially affect task completion time (see Chapter 3).

When navigating using AR, subjects were shown the camera feed overlaid with a 3D model of the skin (opaque), the cortex (translucent), the target (red dot), the tip of the pointer (crosshair) and an extension to the pointer, extending 5 cm downwards from its tip (blue line). This extension was meant as a guide to help subjects determine the best possible insertion angle, prior to inserting. When navigating using the traditional IGNS view, subjects were presented with a 3D model comprising of all the same parts, represented the same way, as in the AR view: skin, cortex, target, pointer tip and extension. Additionally, in this mode the standard cut planes of the MRI scan: coronal, transverse and saggital were displayed, see Figure 26. The target was also shown in the corresponding cut planes, as a red circled cross. When in IGNS mode, subjects could switch to see the 3D models in full screen, thus removing MR slices from the view. While they could not move the 3D model themselves, they could ask the experimenter to rotate or zoom in. This replicates how a surgeon would typically have to ask a technician to interact with the system, once they are sterile.

For each system, subjects completed four trials of varying difficulty, resulting in a total of eight tool insertions. For each trial the target to reach was randomly generated, but placed within a zone corresponding to the difficulty level. The generation zones were toroidal regions directly beneath the craniotomy location. A

Table 2: User study difficulty levels

	Level 1	Level 2	Level 3	Level 4
Target depth	$[1, 4]$ cm	$[1, 4]$ cm	$[4, 7]$ cm	$[4, 7]$ cm
Target angle with normal	$[0, 15]^\circ$	$[15, 30]^\circ$	$[0, 15]^\circ$	$[15, 30]^\circ$

table summarizing the difficulty levels is shown in Table 2 and a view of the four regions is shown in Figure 27. This target generation method was chosen in order to have targets sampling the whole space so as to reduce location bias, while at the same time ensuring a broad range of depths and angles.

5.4.3 Experimental Procedure

Before commencing the study, a power analysis was performed using G*Power 3.1 [31] to assess the sample size needed. Given an expected intermediate effect, a sample size of 17 was used. All subjects were briefed on the study, the task and both systems. They were then shown the working of one of the system set-ups and given as much time as they wanted to practice using it and get comfortable with it. Once they felt ready, they completed the first four trials on that system. Next, they were trained on the second system and completed the four last trials on it. Both the order in which the systems were used, as well as the order in which the levels were performed on each system was randomized. To end each trial, subjects were instructed to tell the experimenter when they thought they had reached the target, at which point the trial was ended by the experimenter. After completion of the study, participants were asked to fill in a questionnaire comprising of demographic, prior experience questions, questions on usability and further feedback. They were also asked to fill out the NASA TLX [48] for each systems. A copy of the questionnaire can be found in section A.3.

5.5 User Study Results

The study was run with 8 males and 9 females, aged between 18 and 67, with a median age of 26. Subjects were novices or intermediately experienced users: 35% had prior experience with using IGNS systems, 76% had prior experience with using augmented reality and 59% had experience with reading scans or looking at anatomical data.

System accuracy measurements were taken during the study. The average camera calibration reprojection error was 0.84 mm ranging between 0.6 mm and 1.02 mm. Registration RMS error was 1.32 mm. Pointer tip calibration was performed using fCal, from the PLUS toolkit [68]. The pointer tip calibration error was between 0.15 mm and 0.26 mm. The latency in augmentation display, which varies depending on the complexity of rendered structures and the resolution of the images sent, was estimated to be 50 ms. This estimate was calculated with images of typical complexity for navigation purposes (800×600 pixels, see right side of Figure 26). The latency was computed by synchronizing both devices to a common NTP server and comparing timestamps.

5.5.1 Quantitative Measurements

All data analysis was performed using MATLAB 2018b (The MathWorks, Inc., Natick, USA). First, to check for potential errors in the data, MATLAB's built-in function for outlier detection was run. This was run independently for all time measurements and all distance to target measurements for each condition. It was found that three individual trials were more than three standard deviations away from the mean for time taken and another trial was more than three standard deviations away from the mean for final distance to target. While we do not know specifically what happened during those trials, we posit that an event occurred during the experiment, resulting

in this unusually large difference. These four individual trials were thus removed from further analysis, but only that relating to the affected variable.

To ensure validity of the acquired results, all time to completion measurements and final distance to target measurements were compared group-wise on the basis of difficulty level, trial number, subject ID and prior experience level. No subject was significantly different from any other; subject performance was within a similar range and thus comparable. Prior experience level did not affect performance significantly. However, since our sample size was not evenly distributed relative to prior experience, it is possible that our sample size was not sufficient to show statistical significance. It was found that the time to completion of the hardest difficulty level was worse than all other trials. Specifically, a one way ANOVA revealed a significant effect ($F(3, 131),\ p < 0.0006$)). This difficulty level was the least realistic, being deep and at a strong angle with the craniotomy normal. For both of these reasons, it was removed from further analysis involving completion time.

Since there was some jitter in the pointer tip position, the final distance to target was computed as the minimum distance between tip position and target in the last two seconds of a trial. The three main outcomes of the validation procedure were time taken to reach target, final distance to target and the results from the NASA TLX. Boxplots of final distance to target for all trials and of total time taken to reach targets are shown in Figure 28. Comparative results of the TLX are shown in Figure 29.

For both measures, it was found that MARIN was statistically lower (better) than the standard IGNS guidance view. Specifically, a one-way repeated measures ANOVA showed that there was a significant effect of system type on the final distance to target ($F(1, 136),\ p < 0.04$). For time taken to reach the target, a one-way repeated measures ANOVA showed that there was a significant effect of system type on the

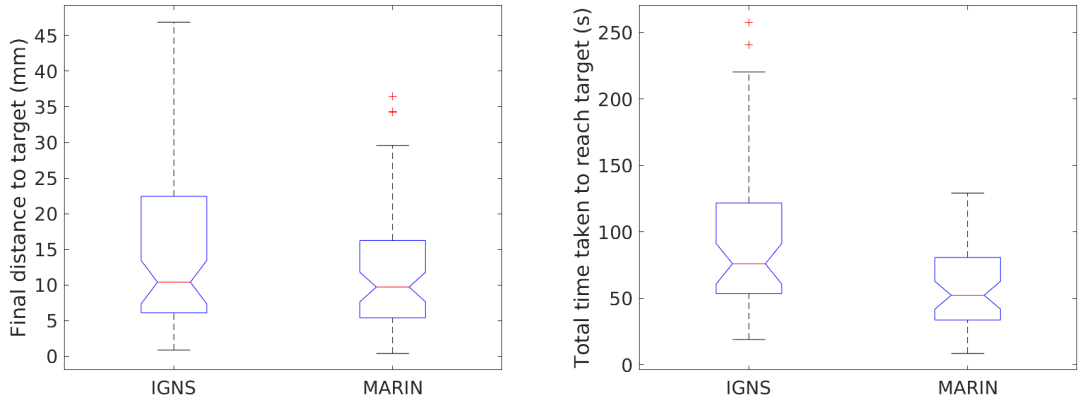

Figure 28: On the left: Boxplots showing distribution of final distance to target, compared per system; On the right: Boxplots showing distribution of total time taken to reach target, compared per system.

completion time ($F(1, 99),\ p < 0.0004$).

Subject's accuracy was found to be uncorrelated to either target depth or target angle from normal. However, it was found that time to completion was weakly correlated with both target depth ($p = 0.026$) and angle with normal ($p = 0.045$).

5.5.2 Qualitative Feedback

In addition to filling out the NASA TLX for both systems, users were also asked which system they found most intuitive to use, which they felt most comfortable using and which they preferred. 100% of subjects found MARIN to be more intuitive to use. 82% felt more comfortable using MARIN, 12% felt more comfortable using the standard IGNS and 6% did not have any preference. 94% preferred MARIN overall, while the remaining 6% did not have any preference.

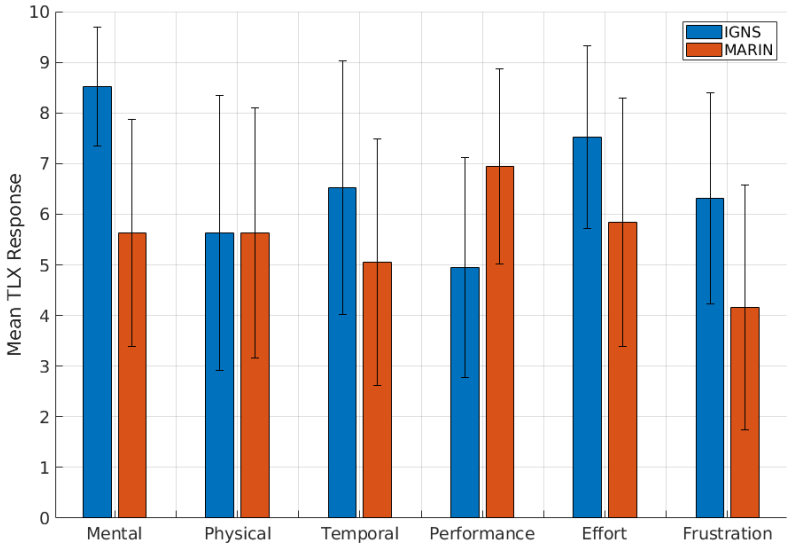

Figure 29: All TLX scales, ranging from 1 to 10 (lower is better on all scales except Performance), compared per system, where error bars correspond to standard deviation.

5.6 Expert Feedback

In addition to the user study, feedback was provided by senior neurosurgeons who tested the system both in the lab, as well as in the OR in clinical cases (Figure 30). Although the system has yet to be rigorously tested in the operating room, we have had continuous feedback from senior neurosurgeons during the development process. Although all found the system to have great potential, opinions differed as to how they would like to use it. Some said they would like it to be clamped to the bed, between them and the operating field and work from above it, using it throughout the operation. Others preferred to use it at the beginning of surgery, looking at the patient from different angles to give them a good sense of the anatomy to better plan their procedure. They would then give it to the nurse when operating to have both hands free, but said they might take it back later on during the procedure for more guidance. They also said they thought it might prove particularly useful for harder cases or those with more intricate anatomy.

Figure 30: Initial tests of MARIN in a clinical setting.

5.7 Discussion

The main advantages of our system over previously published mobile AR systems are improved usability of the prototype, improved responsiveness, increased versatility, customizability, interactivity and wider possibility for adoption. MARIN can run on most hardware and OSes and is entirely open source. It allows users to manipulate data and interact with the AR view through intuitive gestures. The latency in image transfer with MARIN is roughly eight times lower than the lowest latency reported for previously proposed systems. As shown by Sielhorst *et al.* [113], the latency of previous systems severely hindered their usability. Our system on the other hand is in the range they recommend to increase the performance of subjects using it. This gain in performance is attributable to how we transmit augmentation images to the mobile device. Our architecture, similar to that of Watanabe *et al.*'s [128] system, consists of a desktop, a mobile device, a tracking system and a router. To alleviate the problem of delay, MARIN incorporates a modern video encoder, greatly reducing

network load and allowing for shorter latency. Even though time for encoding and decoding is added, the gain in compression much more than compensates for it. The compression ratio depends on the complexity of images sent, but for a typical scene, compression was between 97% and 99%. Even for more complex scenes, this ratio would not be much lower and should remain above 94%, thus still being relevant for all use cases. The remaining latency observed with MARIN is attributable to network transfer, as well as encoding and decoding of the images.

Other IGNS system performance metrics, (*i.e.,* camera calibration reprojection, registration RMS and pointer tip calibration error) place our system in the same range as previously proposed AR navigation setups. In our user study, MARIN outperformed the traditional navigation system, which served as our baseline for comparison, both in terms of time to reach the target, as well as accuracy with which users were able to reach the target. Unfortunately, a direct comparison with other mobile AR systems cannot be made since no similar quantitative user accuracy assessment were done for any of them. Although, considering findings from Sielhorst *et al.* [113], it can be hypothesized that thanks to MARIN's improved performance, users would perform better with it than with previously proposed ones.

Qualitative feedback unequivocally showed that MARIN outperforms standard IGNS views in terms of usability, comfort and ease of use. Subjects agreed that MARIN was more comfortable, easier to use and overall preferred it. The TLX results confirm these findings, as MARIN received better ratings on all scales.

An interesting point is how subjects made use of the two different systems. While navigation using the AR view was more or less constant across the sample, navigation using the standard IGNS system varied wildly. Some subjects looked mostly at the slices for guidance, finding the location of the target in all three planes and mentally mapping the location before making an insertion, while others looked almost only

at the virtual view and tried to match the pointer extension with the target in that view. Thus individual differences in guidance system usage are present depending on background and prior familiarity with certain tools, views or concepts. We believe that this highlights one of the greatest strengths of MARIN: it lets the user decide, through an intuitive interface, what they want to see, enabling them to switch between AR, VR or standard IGNS views and also easily control through gestures what is displayed and how it is displayed in the AR view.

MARIN can integrate seamlessly into the typical surgical workflow. It can easily be draped in a sterile bag in the OR. It is less cumbersome than some other types of AR setups, such as HMDs, which can be bulky and disruptive, as well as not well-adapted for surgery [14]. A mobile device, on the other hand, can be stowed away in a matter of seconds and brought back if needed. Additionally the gestures used for interaction are familiar to anyone who has used a mobile device. Thus interaction is more natural and intuitive than that of other devices such as, for example, the HoloLens™. A mobile device can also be moved freely around the patient, allowing the view of the anatomy from a wide range of angles, thus providing a good sense of depth through the parallax effect. Further, this allows the surgeon to explain the approach and anatomy to residents and students in the operating room.

5.8 Conclusion

Mobile devices offer new untapped potential still to be harnessed in the surgical domain. MARIN is a step in this direction. With the release of MARIN as an open source project, we hope to pave the way for others wishing to build mobile AR setups for surgical guidance. MARIN can be used in conjunction with a broad range of hardware and software making it easy to integrate in an array of systems. The results presented here show that it has strong potential to make guidance more

intuitive, easier and more comfortable to use. The evidence gathered here suggests that mobile AR may also enable shorter operation time and more accurate navigation. Although, to confirm this, quantitative measurements will need to be made with surgeons in a clinical setting. MARIN also enables other potential future research directions, including integrating additional data sources provided by mobile devices (*e.g.,* accelerometer, gyroscope, etc.) to improve tracking accuracy or even explore the possibility of making an affordable and compact IGNS system by relying on mobile device trackers rather than an external tracking system. These are future avenues of research we are working on.

Addendum

In an extension to the initially published system, a video-based hemodynamic feature was added. The developed method allows a user to capture a video sequence with the mobile device (*e.g.,* iPad) intraoperatively, select some vessels on screen using touch gestures and obtain blood flow direction information displayed in the *in situ* augmented reality view. Intraoperative blood flow directionality information is a very valuable asset for surgeons to have access to during neurovascular cases and particularly complex ones. Assessment of the method proved its feasibility and that mobile devices are suitable for such an application. Additionally, this extension of the original work shows how MARIN's versatility and upgradability offers potential for improvements.

A prototype for video-based hemodynamic analysis in neurosurgery implemented on a mobile augmented reality system

5.9 Purpose

Cerebrovascular surgery is the method of choice to treat arteriovenous malformations (AVMs), arteriovenous fistulae (AVFs) and aneurysms. AVMs and AVFs consist of abnormal collections of blood vessels in the brain, while aneurysms are bulges due to weakness in the wall of a blood vessel. Surgical treatment of AVMs and AVFs consists of clipping the feeding artery or arteries prior to tying off the draining veins and removing the nidus. Clipping the wrong vascular branch can result in severe complications such as haemorrhage, which can lead to neurological deficits. It is

thus paramount for neurosurgeons performing neurovascular procedures to be able to distinguish between feeding and draining vessels in and around the pathology. The vessels associated with these vascular malformations often appear as neither true arteries or veins, so practitioners must rely on preoperative imaging to determine blood flow directionality. However, translating this information from preoperative imaging to the surgical scene can be challenging. For this reason, methods such as intraoperative angiography with dye injection, intraoperative Doppler ultrasound acquisition and hemodynamic video analysis have been proposed in the past to provide hemodynamic information. The latter has several advantages, including the fact that it is contactless, contrary to Doppler ultrasound, where external force may damage the already weakened vessel walls. Video analysis can also be performed continuously throughout the surgery, unlike dye injection, which can only be used periodically, since the dye stays in the bloodstream long after injection (more than 15 minutes). Video-based hemodynamic analysis has only been recently proposed [134] and has yet to be tested intraoperatively. The work presented here shows our initial tests of integrating hemodynamic analysis into MARIN, an open source Mobile Augmented Reality Interactive Neuronavigation system.

5.10 Methods

MARIN consists of an iPad tablet and a desktop computer running the IBIS neuronav platform. A video sequence of a few seconds recorded in 1280×720 resolution at 30 frames per second is acquired on the iPad. After acquisition, this sequence is sent to the desktop using the wireless local area network (WLAN). A region of interest (ROI) is then selected by the user. This is done by picking contour of the vessel region, on the touchscreen of the tablet. The points are sent to the desktop as well, over the WLAN, using the OpenIGTLink protocol and library. The video sequence is then

cropped to keep only the central part of the image. 25% of the original image size was thus removed in order to reduce computation times. This is done since the tablet's camera has a much smaller zoom factor than a neurosurgical microscope. The ROI therefore consists only of a smaller region, which should usually be roughly in the center of the frames. The hemodynamic analysis is then performed on the desktop computer. This analysis consists of image stabilization, motion reduction, intensity magnification in the selected frequency range, as well as filtering, as described in [124]. Once the hemodynamic analysis is completed, a frame showing an augmented reality view of the scene with the direction of blood flow in the selected vessel of interest is sent back to the mobile device for *in situ* AR visualization. Tests were performed in the lab using a blood vessel phantom and test setup as in [124]. The setup and resulting AR view are are shown in Figure 31.

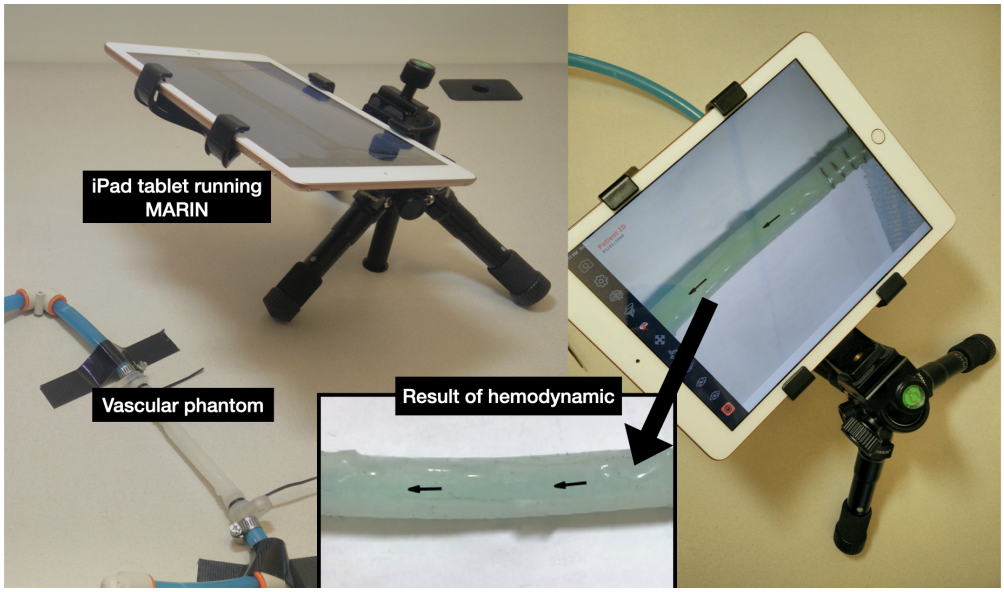

Figure 31: Left: Test setup with the iPad running MARIN on the right and the vascular phantom on the left. Right: tablet running MARIN showing the AR view with hemodynamic results. Insert: Close-up of the hemodynamic analysis results.

5.11 Results

We demonstrated that mobile device's video capture was sufficient for hemodynamic analysis. We believe that the novel ability which allows the surgeon to interact directly with the touchscreen of the iPad to select the vessels of interest, as well as seeing the AR view *in situ* will both be particularly useful, and allow the system to be easy to use and integrate well into the surgical workflow. Interactive selection on the mobile AR neuronavigation app will enable the surgeon to select a region of interest themselves, without having to rely on a technician. This may potentially streamline procedures, since the tablet can easily be used intraoperatively as we have seen in bringing the system into the OR for AR guidance. Our initial evaluations have shown that, since it is a fairly small and wireless device, it can easily and quickly be brought in when needed and stowed away once guidance is not necessary. It can also be draped in a sterile bag to allow its use throughout the procedure, giving full control to the operating surgeon. The current work presents our initial feasibility tests on integrating hemodynamic analysis into a mobile AR neuronavigation system. In future work, we will provide a more seamless integration of hemodynamic analysis into MARIN that will involve allowing for real-time augmentation that is updated as the tablet is moved around the scene.

5.12 Conclusions

The work presented in this abstract shows that the hemodynamic analysis presented in [124] can be integrated in a mobile AR neuronavigation platform. The camera of the tablet device has sufficient frame rate and resolution to estimate correct flow directions in vessels. It also enables easier interaction for the practicing surgeon with the selection of regions of interest for the blood flow computation. The next step of

this work will involve bringing the system into the operating room to determine its functionality intraoperatively during neurovascular cases.

CHAPTER 6

Evaluation of low-cost hardware alternatives for 3D freehand ultrasound reconstruction in image-guided surgery

Preface

In this chapter, we explore the effectiveness of low-cost tracking and intraoperative ultrasound in the context of image-guided neurosurgery (IGNS). With traditional IGNS systems, the tracking of the surgical instruments is done using highly accurate but expensive optical measurement systems. By using a camera or depth sensor rather than a designated tracking system (the most expensive and cumbersome hardware

component of an IGNS system) a more affordable and accessible system can be developed. As well as cost-effective tracking solutions, in recent years portable low-cost ultrasound systems have been emerging on the market. The study presented in this chapter evaluates the feasibility of using these new and more affordable tools in the context of image-guidance. Specifically, we look at using three-dimensional US freehand reconstruction which enables building an intraoperative volume from a sequence of 2D slices acquired with an ultrasound probe. The reconstructed volume can then be registered against preoperative images to correct for movement and deformation of the anatomy (*e.g.,* brain shift) in order to provide more accurate guidance.

In this chapter, we assess reconstruction accuracy of different hardware combinations at different price points, including a recently proposed sensor fusion method running on low-cost hardware. To test the different hardware possibilities a simple phantom and protocol were devised to enable assessment in three dimensions. We found that the lower-cost ultrasound probe and mid-range optical trackers were suitable for this application and worked similarly to the more expensive components. However, the sensor fusion method running on a depth and RGB camera didn't yield sufficient accuracy to reach clinical acceptance. The developed protocol enables future method developments to be tested within a comparable framework. A shortened version of this paper was submitted to the International Workshop of Advances in Simplifying Medical UltraSound (ASMUS).

Abstract

Current commercial image-guided neurosurgery or neuronavigation systems have a high initial cost and require trained technician assistance for accurate registration and set-up making them unavailable in many hospitals and operating theatres. The evolution of consumer-grade hardware components (*e.g.,* trackers, portable ultrasound probes) has opened the door for the development of low cost systems. In this chapter, we evaluated different low-cost tracking alternatives on the accuracy of 3D freehand ultrasound reconstruction in the context of image-guided neurosurgery. Specifically, we compared two low-cost tracking options, an Intel RealSense depth camera setup (less than 200 USD) and the OptiTrack camera (2.3k USD), to a standard commercial infrared optical tracking system, the Atracsys FusionTrack 500 (∼25k USD). In addition to the tracking systems, we investigated the impact of ultrasound imaging on 3D reconstruction. We compared two ultrasound systems: a low-cost handheld ultrasound system at 7k USD and a high-resolution ultrasound mobile station (over 250k USD). Ten acquisitions were made with each tracker and probe pair. Our results showed no statistically significant difference between the two probes and no difference either between and high and low-end optical trackers. However, the sensor fusion RGB/depth based system did not perform well enough to be usable in this scenario. The findings suggest that some lower cost hardware may offer a solution in the operating room or environments where commercial hardware systems are not available without compromising on the accuracy and usability of US image-guidance.

6.1 Introduction

Neurosurgical guidance systems have shown positive impacts on tailoring craniotomies, reducing interventional errors, increasing tumour resection percentages

and improving patient survival rates. However, in spite of advances in both hardware and software, these systems suffer from accuracy degradation as a procedure progresses and the patient to image alignment computed at the beginning of surgery gets invalidated by the movement and deformation of the brain, or *brain shift*. To reregister the patient intraoperatively, updated images can be acquired, using either intraoperative magnetic resonance images (iMRI) (*e.g.,* Clatz *et al.* [17]), intraoperative computed tomography (iCT) (*e.g.,* Riva *et al.* [104]) or intraoperative ultrasound (iUS) (*e.g.,* Reinertsen *et al.* [101]). The latter is much less expensive, has a smaller footprint in the OR, and has shown that using intraoperative US-based registration correction can account for brain shift [123]. Here, in an effort to democratize access to higher quality neurosurgical care, we explore the use of different low-cost hardware in 3D freehand ultrasound reconstruction, which can be used in the context of intraoperative brain shift registration correction or percutaneous, minimally invasive, needle/instrument guidance. This study thus aims at assessing what the minimum requirements are for a system that may be used in the neurosurgical context, understanding how to reduce costs while keeping sufficient guidance accuracy.

Much research has focused on the use of high-end equipment to achieve increasingly high quality of care for patients. However, the expectation for every center to be able to afford the very high-end equipment is not realistic. In order to increase quality of care worldwide, more research is needed on developing low-cost alternatives to high end systems. In 2019, Malham and Wells-Quinn [72] answered the question of what equipment were most worth purchasing for navigation and imaging systems in the context of spine surgery. They formulated recommendations based on different procedural needs, with arguments such as which types of imaging should be used according to which level of accuracy is required, while also considering other factors

like level of associated radiation. Our proposed research, complements this previous study by looking at the equipment required for 3D freehand US reconstruction specifically, as this is an intraoperative imaging means that not only has important benefits for neuronavigation, but also has a cost that is much lower than other options like iCT or iMRI. Alternative systems relying on iUS imaging are currently the focus of active research to increase accessibility to neurosurgical guidance particularly in the developing world where expensive hardware is out of reach [87].

Although more affordable than MRI and CT solutions, the price range of US systems still varies significantly from a low-cost handheld system (∼5-8k USD) to a high-resolution station (between 50k and 250k USD). In addition to the intraoperative imaging modality, another hardware component used in IGNS to perform 3D freehand ultrasound reconstruction is the tracking system. This component, usually an optical tracking system, accounts for the largest portion of the hardware cost of open-source neuronavigation systems. While different tracking solutions exist, (*e.g.,* eletromagnetic sensing, laser reflection, optical stereophotogrammetry, video-based), the impact of the accuracy as well as the context in which such devices can be used has not been thoroughly studied. In this work, we compared the accuracy of ultrasound reconstruction obtained with different hardware setups at a broad range of price points. Hardware components were selected and US reconstructions were computed with each of these setups. For the ultrasound transducer, two options were compared: a ∼250k USD mobile station and a ∼7k USD handheld system. For tracking four options are compared: a ∼25k USD high-end optical tracker, a ∼3k USD lower-end optical tracker, a sensor fusion hybrid tracking method which uses a ∼200 USD depth camera and an camera-based method using a ∼20 USD equivalent RGB camera. Using these different set-ups we aimed to answer the following questions: Can compromises be made on some of the components without sacrificing too much

on accuracy of the 3D freehand US reconstruction? If so which ones and how is a given budget best invested between these components?

6.2 Previous Work

Cenni *et al.* [15] looked at the effect of using different hardware setups on 3D freehand US reconstruction quality. The focus of their work was not on the influence of hardware or hardware cost on achievable accuracy but rather on evaluating a novel reconstruction algorithm. They tested their method with two different optical tracking systems (exact models not disclosed) and report having found no noticeable difference in the reconstruction quality with the two systems. However, in their work, separate acquisitions were made independently with each tracker, making the comparison less robust. The study done here complements this previous work, with a focus explicitly on hardware differences across a wide range of equipment price points. A key difference of our approach relative to Cenni *et al.*'s, that allows for a more direct comparison of the acquired data, is that all tracking information is acquired simultaneously. As well, our study investigates the use of different US probes in addition to the tracking systems for a more complete understanding of interactions between each components' accuracy on the overall results. Finally, our phantom construction method detailed below is more precise than that of Cenni *et al.*'s study, in which wires were used and measured *a posteriori* with a caliper. In our work, we also assess error in all three dimensions of space, instead of only in a plane as done in Cenni *et al.*

In addition to the two commercial optical tracking systems tested in our study, a recently proposed experimental tracking method is also investigated. The low cost alternative tracking method tested in our study is similar to that presented by Asselin *et al.* [2]. In their work, Asselin *et al.* developed a sensor fusion tracking method that

uses a depth camera and an RGB camera to detect an ArUco marker in the RGB image to determine the x and y position in space and the depth camera to determine its z position. They found that the method worked very well, much better than when using an ArUco marker alone and would be suitable for intraoperative tool tracking. The present study extends their work in assessing if that low-cost hardware and method can be used for 3D freehand reconstruction. The last tracking method tested is that of using an ArUco marker alone. It serves as our baseline, similarly to Asselin *et al.*'s study.

Our work also complements the work done by Gueziri *et al.* [46] which investigated the influence of different parameters on 3D freehand US reconstruction quality. In their study, Gueziri *et al.* investigated the effect of probe depth and frequency on the resulting reconstructions. They found no statistically significant difference between the registration accuracy for all depth and frequency values tested.

In terms of ultrasound reconstruction, a number of studies have investigated the reconstruction quality of different 3D freehand reconstruction algorithms. The interested reader is referred to Solberg *et al.* [114] and Rohling *et al.* [108] for a review in this area. A review of US probe calibration in the context of 3D freehand US reconstruction is also available in [80].

The contributions of this work are: (1) devising a reproducible phantom and protocol to evaluate 3D freehand US reconstruction distortion in three dimensions; (2) investigating the effect of more variables on reconstruction quality and their relations within that robust and tailored experimental protocol; (3) evaluating feasibility of using a cheap camera sensor fusion method for this application; and (4) evaluating impact of reconstruction on subsequent usage of volume in a typical surgical guidance scenario (volume-based registration).

6.3 Methodology

In order to make a fair assessment of a system's cost on guidance performance and to evaluate where component cost has the strongest impact on overall reconstruction accuracy, we designed an experiment to simultaneously test both the imaging modality and the tracking modality.

6.3.1 Hardware Setups

The acquisitions were made using two ultrasound probes in combination with four tracking systems. The tested ultrasound probes are: (1) the MicrUs MC42R20S3 probe (TELEMED, Vilnius, Lithuania); and (2) the BK3500 14L3 probe (BK Medical, Peabody, MA, USA). The tracking systems that were tested are: (1) ArUco markers [36] captured with the RealSense RGB camera; (2) the RealSense D435 (Intel Corporation, Santa Clara, CA, USA); (3) the Optitrack V120:Duo (NaturalPoint Inc., Corvallis, OR, USA); and (4) the FusionTrack 500 (Atracsys LLC, Puidoux, Switzerland). All combinations of ultrasound and tracker were used to capture ultrasound acquisitions. Figure 32 shows the different devices used for acquisition.

6.3.2 Phantoms and Marker Construction

To enable the most precise and fair assessment possible, a phantom and markers were designed for the experiment. Similarly to that of Cenni *et al.*, a wire phantom was constructed. The phantom, built from Lego™, has eight wires pulled taughtly between Lego bricks. The wires thus form a cuboidal shape of precisely known dimensions. All wires cross perpendicularly. Angles between line segments are thus also precisely known. Lego bricks themselves are accurate to within 0.04 mm [71] and the wires were carefully pulled between them, which translates to a very accurate

Figure 32: Depiction of trackers and probes used for acquisition. Top-Left: Atracsys FusionTrack 500; Middle-Left: Optitrack V120:Duo; Bottom-Left: Intel RealSense D435; Center: Telemed MicrUs; Right: BK 3500. Note: devices are not to scale.

phantom. A picture of the phantom is shown in Figure 33. The cuboid measures 11.20 mm by 9.60 mm by 19.00 mm in x, y and z respectively. The phantom was immersed in water for US acquisition.

A custom marker, similar to that of Asselin *et al.* [2], was designed to enable all tracking methods (*i.e.,* both RGB camera and optical) to capture the position of the probe in the same coordinate frame. The marker pivot (3D position) for all tracking methods was defined to be a common point at the center of the construction (the center of the ArUco marker, which corresponds to the centroid of the reflective sphere positions). The marker was 3D printed on a Raise 3D Pro2 printer (Raise 3D Technologies, Inc., Irvine, CA, USA) using a 0.1 mm layer height. A rigid probe attachment bracket was also designed and printed with the same printer settings. A picture of the marker attached to the probe is shown in Figure 34. A similar marker was designed as a reference and attached to the phantom. This custom design and

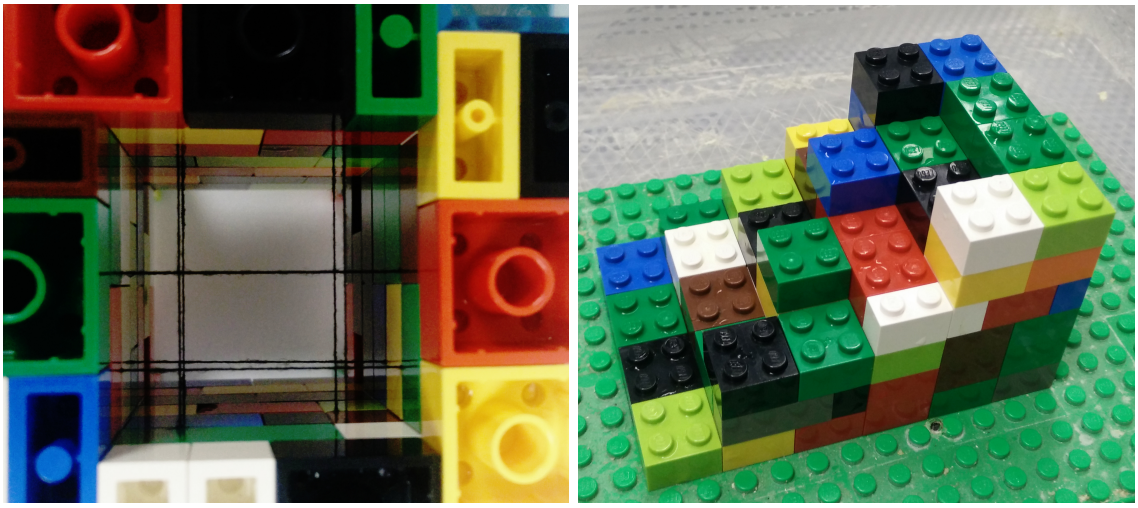

Figure 33: Left: Picture of the experimental wire phantom. Right: Picture of the phantom from [135].

precise alignment of the tracked position for all trackers was done in order to reduce potential bias in the comparison.

6.3.3 Experiments

The experiment was designed to minimize potential bias and ensure a true comparison between systems. Tracking was captured simultaneously with all tracking systems for each US acquisition. All trackers and probes were set to capture data at 30 frames per second to remove frame rate as a potential confound. Trackers were each placed at their optimal working distance from the scene to mimic what would be done in a real-world scenario. Trackers were also each placed as close as possible to the same viewing angle with respect to the scene, again trying to minimize measurement volume as a confounding factor. The trackers were all aligned with the phantom so that the axes of the tracking volume would match that of the phantom. See Figure 34 for a picture of the tracker arrangement during acquisition. To simplify the setup the live camera feed used for detecting the ArUco markers was that of the RealSense. This enabled

us to have only three physical devices in the test setup while allowing testing with four tracking methods. The resolution of the RealSense RGB camera is 848 by 480 pixels, which is low for modern hardware. So, even though it was captured on a more expensive device, it could be achieved just as well with a $20 webcam.

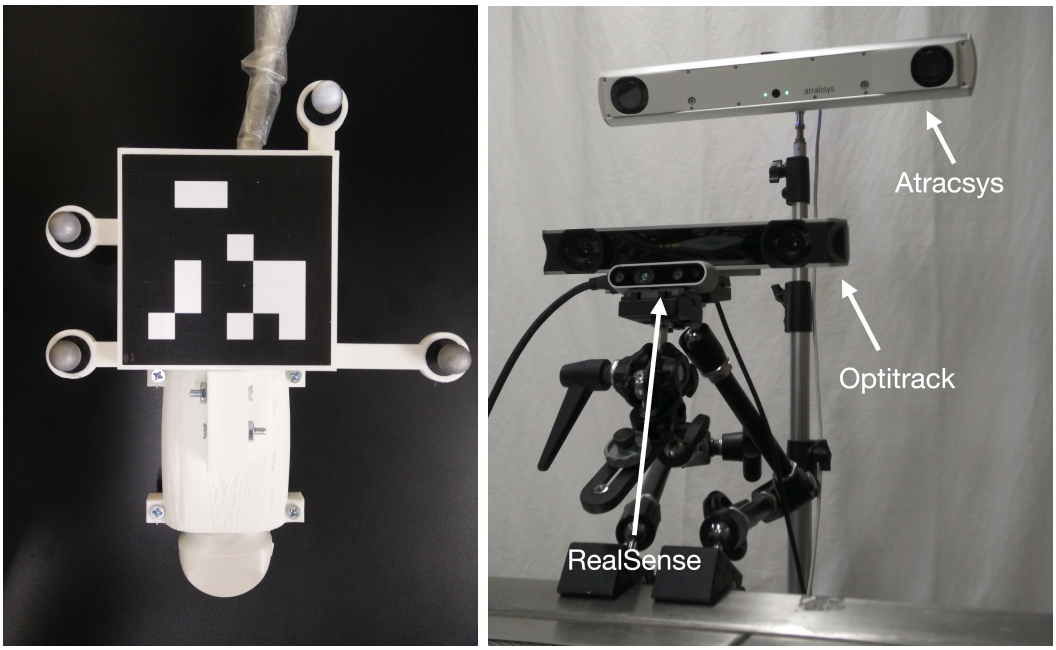

Figure 34: Left: Picture of the custom hybrid markers and probe attachment. Right: Picture of tracker arrangement for data acquisition.

The ultrasound probes were calibrated, both temporally and spatially, using fCal from the PLUS (Public software Library for UltraSound imaging research) toolkit [68], version 2.8. fCal implements the 3 N-wires calibration procedure [13], a method that was previously shown to be reliable and accurate [79]. This calibration was computed for each ultrasound probe with the tracker corresponding to the high-end of its price bracket (for the BK imaging system, the FusionTrack 500 tracker and for the Telemed imaging system, the V120:Duo tracker). The reasoning behind this was that using similarly priced devices in a system might be a more common use case.

Two sets of acquisitions were done. The first one was designed to measure the

distortion in the shape of the reconstructed volume, using the phantom described in the previous section. The second one was designed to test the performance of the reconstructions in its intended use case, intraoperative volume-based registration. For that second experiment, the Lego phantom described by Xiao *et al.* in [135] was used, see right side of Figure 33.

US Reconstruction Accuracy Experiment

For the first set of acquisitions, ten sweeps of the phantom were acquired with each ultrasound probe. For each acquisition, the tracking data was recorded simultaneously with all tracker systems. All sweeps were done in one linear motion done along the z axis. Independent reconstructions were then computed from each sweep and hardware combination, using the PLUS reconstruction [68].Thus, from the 20 acquired sweeps, a total of 80 volume reconstructions were computed and further analyzed for distortion.

On each of these reconstructed volumes, the eight lines corresponding to all wires were manually segmented using 3D Slicer [97]. Figure 35 shows a sample segmentation of a wire phantom acquisition. The order in which segmentations were performed for each trial was randomized between conditions to reduce potential operator bias. All segmentations were performed by the same operator. The intersection of the eight line pairs corresponding to the corners of the reconstructed cuboid were computed in a least-square sense. These eight constructed points were then used in all further analysis. The distance between these points were used to compute the dimensions of the cuboid, or the dimensional distortion (DD) along each axis and angles between the line segments were used to compute angular distortion (AD) around each axis. All metrics were averaged over each axis. This means that for the dimensions, all four segments spawning from the connections of points along that axis are averaged. As well, for angles, all eight angles corresponding to rotation around that axis are

averaged. This averaging gives more robustness to the protocol and reduces the effect of uncertainty associated with segmentation.

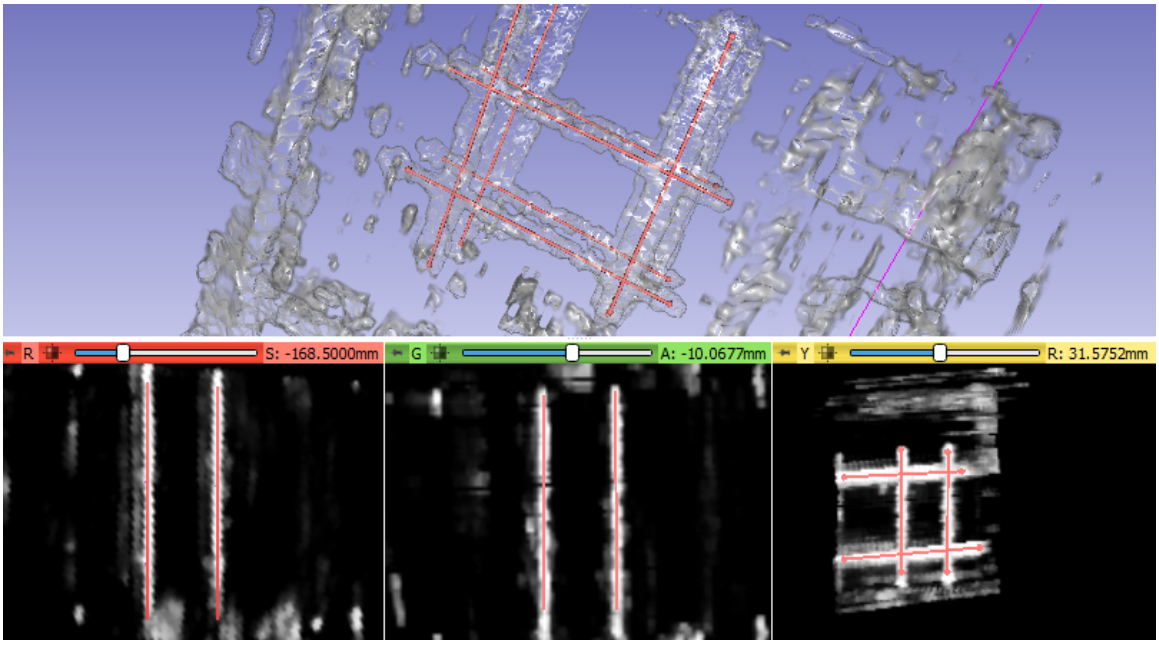

Figure 35: Example wire phantom segmentation. The segmented wires appear as red lines in both the slices views and the 3D view.

Preoperative Volume Registration Experiment

For the second set of acquisitions, using the phantom from [135], a very similar process was followed. One acquisition was made with each probe where all tracking data was captured simultaneously. In order to capture the whole phantom, sweeps were done in two parallel motions done along the z axis. Independent reconstructions were then computed for each hardware combination. The eight resulting volumes were then all registered independently to a model of the phantom. For each reconstruction, the registration process was as follows. The model and reconstruction were loaded into 3D slicer [97]. The preoperative model was manually brought into rough alignment with the reconstruction. The volume registration module of 3D Slicer was used to

compute the registration transform with the Mattes Mutual Information (MMI) cost metric [75]. The RMS target registration error (TRE) was computed using seven target points distributed across the phantom, that were easily identifiable in both the model and the volume (center of cylindrical protrusion on Lego blocks).

6.4 Results

To determine the shape distortion of the reconstructed volume we calculate the dimensional distortion (DD, or discrepancy in length of the reconstructed object) along all dimensions and the angular distortion (AD or discrepancy in angles between substructures) along each axis. For both DD and AD, the absolute value of the error is used in analysis as both a negative or positive error would have similarly undesirable effects on the usability of the resulting reconstruction. DD is reported as a percentage error of the supposed length value and AD is reported as an angle difference from the supposed angle (90°). In addition to error assessment along all axes independently, a compounded percentage volumetric error was also computed. This error metric, while less telling than the independent per axis one, enables easier comparison with previous work, as it is usually reported in previous literature. Both DD and AD are averaged over each axis, as this averaging enables smoothing noise that could have arisen from the manual segmentation.

The image to probe calibration reprojection error for the Telemed system was 0.87 mm and for the BK system 1.26 mm. The temporal calibration yielded a 38 ms latency for the Telemed and a 48 ms one for the BK.

6.4.1 US Reconstruction Accuracy Results

After having performed the segmentation on all 80 volumes, an outlier detection procedure was performed. This consisted in looking at all metrics on a per setup basis (combination of tracker and probe) and flagging any individual segmentation where some values were more than two standard deviations away from the mean. This could indicate a potential mistake in the segmentation of one or more of the wires. The flagged segmentations were rechecked and corrected when needed. This enabled us to minimize the amount of manual errors arising from segmentation.

Table 3 shows the dimensional distortion results for all combinations of hardware. Table 4 shows the angular distortion results for all combination of hardware. Results in both tables show the mean value for each setup plus/minus the standard deviation. Figure 36 presents box plots shown per axis and per tracker.

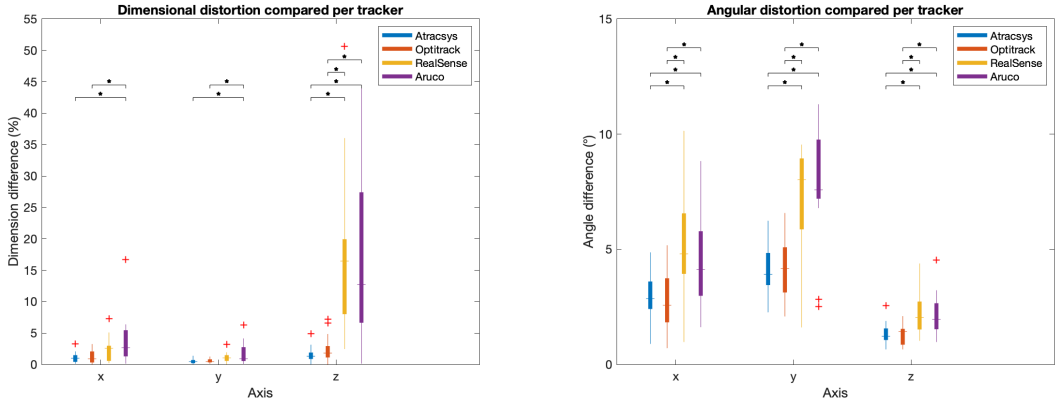

Figure 36: Boxplots of dimensional distortions compared per tracker (on the left) and angular distortions compared per tracker (on the right). Relationships marked with a star ($\star$) are those where the difference between group means are statistically significant to within $p < 0.05$.

We found that the probe used had little impact on the overall accuracy of the reconstruction. A two-way ANOVA revealed that the BK and Telemed reconstructions were not statistically significantly different from one another on

Table 3: Mean dimensional distortion per axis for each combination of hardware. DD values are expressed as percentages of the expected dimensions. Numbers are reported with the corresponding standard deviations for each group. Volume values are expressed as percentages of the expected volume. Devices are ordered in decreasing order of cost.

Probe	Tracker	Dimensional distortion (%)			Volume (%)
		x-axis	y-axis	z-axis	
BK	Atracsys	1.37 ± 0.96	0.59 ± 0.33	1.77 ± 1.29	2.54 ± 2.16
	Optitrack	1.40 ± 1.05	0.66 ± 0.38	2.13 ± 2.16	3.30 ± 2.77
	RS	2.79 ± 2.32	0.82 ± 0.43	12.28 ± 6.53	12.75 ± 9.27
	ArUco	5.06 ± 4.80	1.90 ± 2.16	19.02 ± 10.63	20.73 ± 14.94
Telemed	Atracsys	0.78 ± 0.37	0.41 ± 0.39	1.28 ± 0.85	1.14 ± 0.89
	Optitrack	1.08 ± 1.09	0.47 ± 0.28	2.70 ± 1.80	3.44 ± 2.71
	RS	1.85 ± 1.22	1.36 ± 0.86	21.65 ± 13.72	22.53 ± 13.23
	ArUco	2.64 ± 1.96	1.50 ± 1.20	13.74 ± 13.87	13.83 ± 14.66

Table 4: Mean angular distortion per axis for each combination of hardware. AD values are expressed in degrees. Numbers are reported with the corresponding standard deviations for each group. Devices are ordered in decreasing order of cost.

Probe	Tracker	Angular distortion (°)		
		x-axis (pitch)	y-axis (yaw)	z-axis (roll)
BK	Atracsys	2.79 ± 0.95	3.31 ± 0.56	1.36 ± 0.57
	Optitrack	3.24 ± 1.16	3.36 ± 0.96	1.31 ± 0.48
	RS	5.24 ± 1.38	7.75 ± 2.44	2.29 ± 0.94
	ArUco	5.03 ± 2.12	9.65 ± 1.37	2.16 ± 0.65
Telemed	Atracsys	3.12 ± 1.17	4.82 ± 0.70	1.19 ± 0.35
	Optitrack	2.30 ± 1.31	5.00 ± 1.12	1.36 ± 0.44
	RS	5.00 ± 2.98	6.30 ± 2.38	2.13 ± 0.90
	ArUco	4.26 ± 1.87	6.53 ± 2.18	2.16 ± 1.09

neither dimensional nor angular error on almost any axis. They were only different in the x dimension, where the BK was worse than the Telemed ($p = 0.0275$), for all other metrics they were not statistically different. For that reason, data for both probes was bundled in Figure 36. Reconstructions done with the Atracsys and Optitrack trackers were also not significantly different from one another on any metric and any dimension. All differences that were statistically different from the null hypothesis are:

- Aruco was statistically different (worse) from Optitrack and Atracsys dimensionally in all x and y and z (respectively in x, y and z for Atracsys: $p < 0.001$, $p < 0.001$, $p < 1e^{-5}$; and for Optitrack: $p < 0.0023$, $p < 0.0020$, $p < 4e^{-6}$).

- RealSense was statistically different (worse) from Optitrack and Atracsys dimensionally in z ($p < 1e^{-6}$ for both).

- Aruco was statistically different (worse) from Optitrack and Atracsys angularly in all x and y and z (respectively in x, y and z for Atracsys: $p < 0.017$, $p < 5e^{-8}$, $p < 0.001$; and for Optitrack: $p < 0.006$, $p < 1e^{-7}$, $p < 0.002$).

- RealSense was statistically different (worse) from Optitrack and Atracsys angularly in all x and y and z (respectively in x, y and z for Atracsys: $p < 0.001$, $p < 4e^{-5}$, $p < 5e^{-4}$; and for Optitrack: $p < 3e^{-4}$, $p < 8e^{-5}$, $p < 0.001$).

6.4.2 Qualitative Results

When visually inspecting the 3D ultrasound reconstructions, a number of things can be seen. First, the Telemed and BK are clearly different in terms of image quality. The wires appear more fuzzy in the images acquired with the Telemed. Second, there was a noticeable visual difference in reconstruction quality between volumes obtained with either the Atracsys or Optitrack and those obtained with ArUco or RealSense. In those from the ArUco alone or RealSense the wires are much less clearly defined (those of the ArUco alone being slightly worse). This lower visual quality of the reconstruction translated very strongly to greater difficulty when doing manual segmentation. The ArUco and RealSense acquired volumes were much noisier and jagged, which made the segmentation process more error prone. Picking the center of those wires was harder on those reconstructions than those obtained with the other two systems.

Figures 38 and 37 show sample reconstruction results.

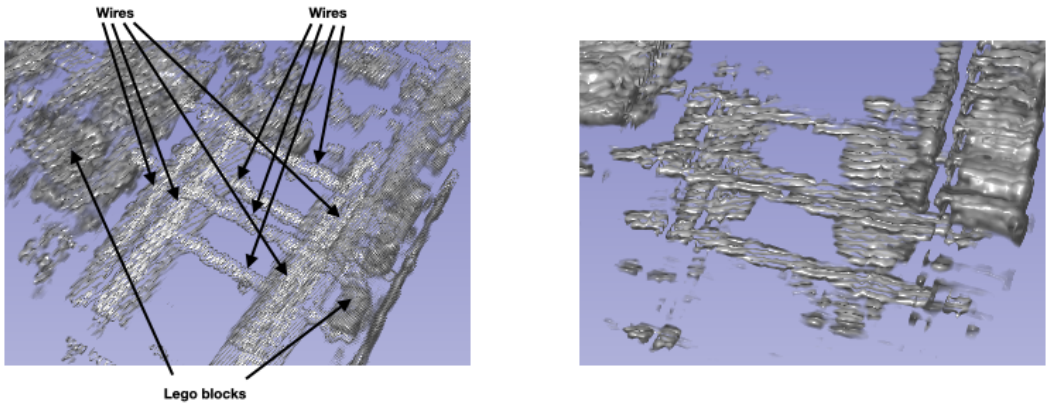

Figure 37: Side-by-side comparison of typical reconstruction results obtained with each US acquisition systems. Both acquisitions depicted were acquired with the Atracsys tracking system. Left: BK imaging system; Right: Telemed imaging system.

6.4.3 Preoperative Volume Registration Results

For the second set of acquisitions done on the Lego phantom, the RMS TRE after registration obtained are those reported in Table 5. As we can see, the results from this second experiment are in good agreement with those from the first one. Both optical trackers performed well, although in this case the low-cost probe produced less accurate registrations than that from the high-end one.

Table 5: TRE after registration for each combination of hardware.

Probe	Tracker	TRE (mm)			
		x-axis	y-axis	z-axis	RMS
BK	FusionTrack 500	1.93	1.11	0.75	2.35
	V120:Duo	0.80	0.43	0.95	1.31
	ArUco + depth	1.15	1.85	1.90	2.89
	ArUco	1.81	5.64	15.81	16.88
Telemed	FusionTrack 500	1.53	1.01	1.74	2.53
	V120:Duo	2.10	2.91	3.83	5.25
	ArUco + depth	2.13	3.29	5.57	6.82
	ArUco	1.89	1.73	6.46	6.95

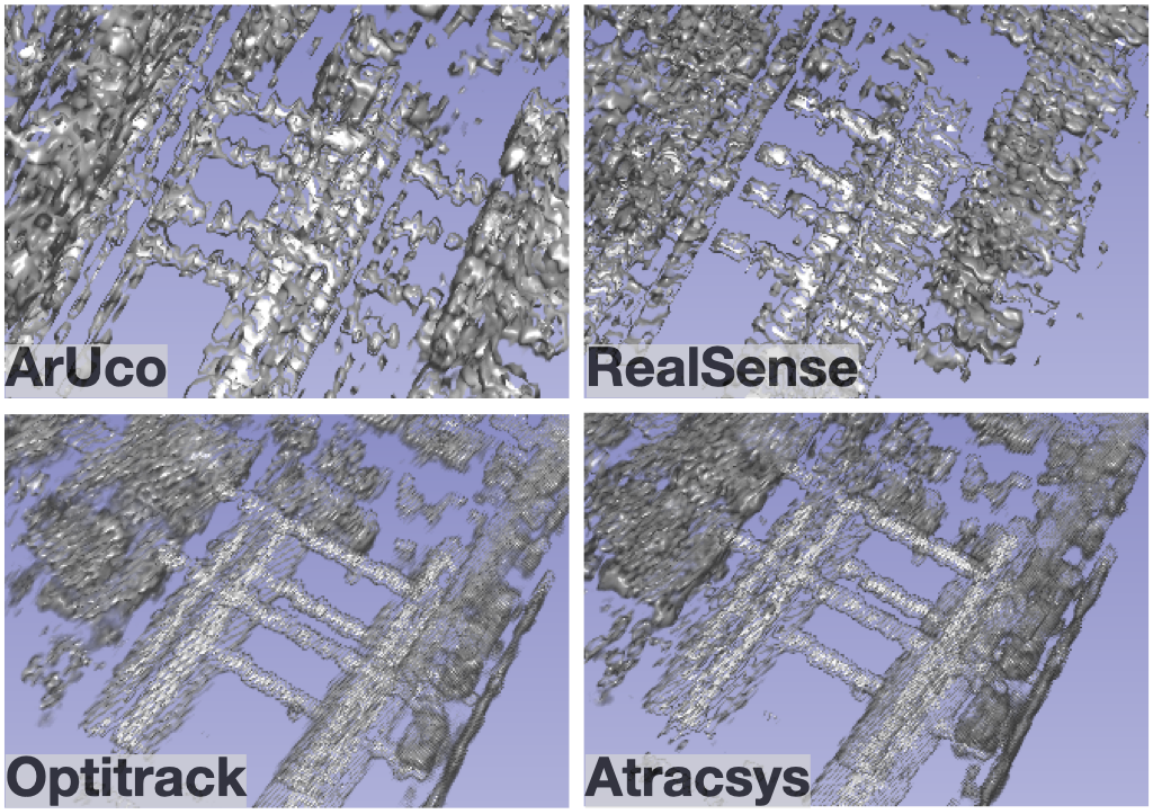

Figure 38: Side-by-side comparison of typical reconstruction results obtained with all four tracking methods. All reconstructions are from the same acquisition, done with the BK imaging system. Top left: Aruco tracking; Top right: RealSense tracking; Bottom left: Optitrack tracking; Bottom right: Atracsys tracking.

6.5 Discussion

The fact that reconstructions made with the BK and the Telemed probe were visually different is not unexpected, the resolution of the BK probe is significantly higher, 728×892 and 12 MHz compared to 512×512 and 4MHz for the Telemed. What is perhaps more surprising is that this apparent difference did not translate into a measurable difference in reconstruction volume quality, meaning that the wires appeared more diffuse but their position corresponded. This finding indicates that there might not be a need for spending more on a higher-end US system for this

application, although training to understand the image based on quality is important. Furthermore, this finding confirms the results from Gueziri *et al.*who found no effect of probe frequency on final registration accuracy. Even though images are less sharp with the cheaper probe, the reconstructed volume is still accurate, which leads us to believe it would perform just as well for intraoperative registration correction or for visualizing tool (*e.g.,* catheter, ventricular drain or needle trajectories), for instance.

The fact that lower-cost hardware (both low-cost probe and mid-range optical tracker) works as well as more expensive hardware hints that error arising from other sources are higher than that of the measurements for all devices, even the cheaper ones. Those other sources are for instance: calibration, reconstruction, unevenness in the sweep acquisition movement, etc.

Something else to consider when comparing US systems is the latency in image transfer. In our experiment and in general in 3D freehand US reconstruction, the time difference between image timestamps and tracking timestamps is assumed to be fixed. The temporal calibration done prior to acquisition enables computing this time difference, which can then be compensated for in software upon data streams arrival. However, it was observed that the BK latency in image transfer was not fixed but rather fluctuated over time. For this reason, the BK was perhaps at a bit of a disadvantage. In our particular setup this fluctuation could have arisen from many sources: US system software, network card drivers, operating system or other receiving software. Although, this point conveys information of general use in 3D freehand US reconstruction, not only about this specific setup. Latency should be considered with great care in this application and efforts should be made to ensure that the latency is not only as low as possible, but also, and very importantly, that it remains as constant as possible throughout an acquisition.

The difficulty described in the previous section in doing the manual segmentation

on the cheapest tracking hardware has consequences beyond just the segmentation process itself. Not only is the process more time consuming for these acquisitions as the viewer takes longer to understand the US images but more importantly this leads to less accurate segmentation. The jaggedness of the reconstructed edges might further impact a registration algorithm as the noise introduces artificial gradients in directions where none should be observed thus gradient-based methods might not work effectively. This will be explored in future work. This less accurate segmentation might be what causes both ArUco and RealSense to be indistinguishable statistically. Values are quite different for the z-axis, but the standard deviation on both samples is also large. There is a possibility that a genuine statistical difference between the two might be obfuscated by this segmentation difficulty due to low quality of the reconstructed volumes.

All systems performed significantly worse in z than they did in the other two directions. This was expected, as all tracking methods tested, be it the commercial optical trackers or the experimental sensor-fusion method, are vision-based, meaning that they measure distances in images. They are therefore always going to be more accurate in the image plane than in the direction perpendicular to it. Although, and while all systems suffered from this, the marker-based (ArUco and RealSense) were much more affected.

The second experiment on the Lego phantom confirms the findings of our first experiment. All conclusions regarding the quality of the reconstructions obtained using different setups translate directly to the accuracy of the registration. All tracking methods and hardware have more error in the view direction. Optical trackers both performed well and better than the methods using image markers alone or in combination with depth. In this case though, the reconstructions made with the cheaper probe seemed to be less accurate. This could however be due to the manual

target segmentation that was harder to do on the lower resolution and blurrier images of the lower cost probe.

Finally, it is worth noting an important limitation in the design of our experiment. Manual wire segmentation, as performed in the experiment, allowed us to compensate for discontinuous data, especially when the quality of the reconstructed volume was low. Although this approach allows for capturing of the overall dimensional and angular distortions, local artifacts such as deformations and mis-reconstructions were attenuated. The effect of these artifacts on the outcome of IGNS applications needs further investigation.

6.6 Conclusion and Future Work

In this study, a phantom and protocol to measure 3D freehand US reconstruction distortion was presented. The wire and Lego phantom is easy and cheap to build and the protocol is easy to replicate. This allows for a more standardized comparison of reconstruction methods and tracking methods in the future. The protocol was used in a study to gain insight into the impact of different hardware components' cost on reconstruction accuracy. Four tracking systems were compared, whose cost were an order of magnitude apart from one another, thus covering a very broad range of price points. As well, two US imaging systems were compared that were roughly two orders of magnitude apart in price. We found that the lower cost US imaging system did not yield reconstructions that were measurably worse than the high-end system. As well, for tracking it was found that the lower-end optical tracking system performed statistically the same as the high-end optical tracker. However, the cheapest sensor fusion tracking methods performed significantly worse. It should be noted that the method by Zhou *et al.* [138] who used the depth camera directly to track the shape of the marker in space, (but was not published at the time of our experiment running)

should also be tested with the same protocol in future work. Seeing as it tracks natively in 3D space, it may perform better than the sensor fusion method tested here. In future work as well, a second series of experiments will be performed to test more specifically how volume registration is impacted by the varying quality of reconstructions obtained with different hardware components.

CHAPTER 7

Conclusion and future work

Improving intuitiveness and interactivity of guidance information during surgery has the potential to improve surgical workflows and in turn patient outcomes. By taking advantage of affordable hardware devices and developing algorithms to address the shortcomings of current systems in terms of accuracy, we can address the need for accurate and affordable neurosurgical care. This thesis explores neurosurgical guidance on both of these fronts, first the use of novel hardware and visualization methods in the context of neurosurgery to both improve intuitiveness and ease of use of current systems, as well as provide lower cost alternatives for deployment in limited resource contexts. Lowering whole system cost can potentially have a significant impact on global health as needs in certain regions are high and mostly unmet at the moment.

7.1 Summary of Findings

This thesis explored integration of new software methods and low-cost hardware to neurosurgical guidance. In doing so, novel interaction methods were devised, a novel system was built and tested and a novel experimental testing protocol was developed. In addition to these contributions, a number of research questions were answered. Table 7 summarizes all research questions explored in this thesis and their conclusions.

7.1.1 Augmented Reality in Neurosurgery: Cognitive Load

In Chapter 3, we explored the effect of augmented reality on performance for neurosurgical task completion. Specifically, we showed that AR significantly reduces the number of focus shifts made by the user while completing a task such as tumour outlining prior to a craniotomy, *i.e.,* removal of the bone flap to access the brain. Reduction of focus shifts, in turn, reduces disruption to the user's task which has been shown to be detrimental to both cognitive and motor tasks. Additionally to reduce attention splitting, which decreases cognitive load, we also found that AR enabled the task to be completed more quickly. This can be a meaningful benefit in the surgical context. Finally, AR presented *in situ* in particular was found to be even more intuitive to use and overall preferred over AR presented outside the surgical area.

Making assessments on the long term effects of using AR technology on either the surgeon, surgical procedure, or patient outcomes requires further study and there is still a relatively small amount of data available on the impact of AR in the operating room. Further, the high variability between surgical procedures means that proportionally more data is necessary to come to meaningful statistical conclusions regarding impact of AR on patient outcomes. Although, what can already be

Research question	Finding
Can AR guidance provided *in situ* streamline and simplify surgical tasks?	Yes. All AR guidance, whether presented *in* or *ex situ*, were found to make task completion shorter and necessitate less attention shifting from the user, which leaves them more focused on the task at hand. *In situ* AR in particular was found to be more intuitive and easier to use by users.
Can touchscreen gestures be used for intraoperative manual brain shift correction?	Yes. The developed method was found to allow users to correct for most of the registration error in a very short amount of time. Users also found the method to be simple and intuitive, making it a good candidate for intraoperative use, offering minimal disruption to normal surgical workflow.
Can AR guidance presented *in situ* enable users to perform a surgical task more accurately?	Yes. Thanks to higher performances of our proposed system relative to previous ones, subjects were able to reach a target within the brain more accurately with it than they were with traditional guidance.
Can mobile devices be used intraoperatively to perform video-based hemodynamic analysis?	Yes, they can. Video sequences captured on a mobile device were found to be suitable for hemodynamic analysis. Integration of such feature to our system enables intraoperative blood flow directionality guidance to be provided *in situ* to operating surgeons.
Can novel low-cost methods and hardware achieve high enough accuracy for clinical relevance in 3D freehand ultrasound reconstruction?	Yes, in some cases. The distortion of the reconstructed volume was found to be too high to be usable for intraoperative registration correction when using the lowest-end hardware and method. However, the low-end ultrasound probe and mid-range optical tracking systems were found to provide results on par with the gold standard, hinting that system cost could be lowered considerably without significantly impacting accuracy.

Table 7: Summary of research questions tested in this dissertation along with obtained experimental results.

assessed is how AR navigation compares to traditional systems in terms of usability, intuitiveness and integration in a typical workflow.

The significance of the work done in this chapter extends previous knowledge of how AR could improve surgical practice specifically from an ergonomics and workflow point of view.

7.1.2 Gesture-based Registration Correction

In Chapter 4, we explored how novel interaction paradigms, enabled by mobile devices, could be used within the context of neuronavigation. A method to correct patient-to-image registration intraoperatively was developed and tested. The method uses the augmented window visualization coupled with gestures registered on a mobile device to perform registration correction. The necessity to correct the registration intraoperatively arises from a number of factors that lead to accuracy degradation during IGNS, *e.g.,* registration errors or brain shift. Brain shift causes the alignment between the patient and the preoperative images that is done at surgical onset to lose accuracy as the procedure progresses. Without an accurate registration between patient and images, the whole navigation system loses its effectiveness and worse, it can lead to errors, which in neurosurgery can often prove highly consequential. The proposed method enables the user through pinch and pan gestures to realign the virtual models to the patient. The process can be done by the neurosurgeon themselves, without the need for a technician to be present. This is a substantial advantage in the surgical context as it has the potential to streamline and shorten procedures, factors that have been proven to impact surgical outcomes.

The method was tested in a laboratory setting to make a precise quantitative assessment of its performance. It was found to be both efficient and work well for medium to large shifts. Users were able to correct most of the error with it in those

cases. On average, users were able to complete the correction procedure in less than a minute. This indicates that the method, in addition to having potential by itself, could also be combined with automated, perhaps non-rigid, methods to make them more robust. It would allow the user to make an initial alignment correction very quickly, that a second method could then refine. This would reduce the risk for the automated method converging on the wrong solution due to a poor initialization, which is a problem that many registration methods are sensitive to.

7.1.3 Augmented Reality in Neurosurgery: Accuracy

In Chapter 5, we explored how *in situ* tablet-based augmented reality guidance could improve users' accuracy in surgical task completion. A complete state-of-the-art neuronavigation system, MARIN (for Mobile Augmented Reality Interactive Neuronavigator), for AR surgical guidance was developed and tested. MARIN is used to present an AR view of the surgical scene using the augmented window paradigm, showing anatomical structures, segmented from preoperative scans, beneath the surface of the brain. The developed system, even though similar in architecture, improves upon previous state-of-the-art systems' performance. The addition of a video encoding and decoding step enabled a significant reduction in network loads in the system, which allowed for smaller latency to be achieved between user movement and augmentation view update. It was shown by Sielhorst *et al.* [113] that even the smallest temporal discrepancies between the two sources of information, real and virtual, have a strong impact on the system's usability and the user's performance when using the system. This is most likely due to the fact that, as was also shown by Holloway [49], a temporal discrepancy of even as little as a few tens of milliseconds causes the perception of a spatial misregistration between the two image sources to be greater than all other misregistration factors. Thanks to video compression, our

system achieved a latency eight times lower than that reported for previous systems.

With this improved *in situ* AR guidance, we were able to compare accuracy of users performing a precise insertion to a target in the brain using AR and using a traditional guidance system. We found AR guidance to be significantly better for this typical surgical task. Users were able to reach a target both more quickly and with higher precision with AR than with traditional IGNS guidance. This finding, along with that of Chapter 3 make a strong case for integrating *in situ* AR guidance into surgical workflows, seeing how it has the potential to improve quality of care. Furthermore, the system was brought into the OR for a few surgical cases and the initial feedback from surgeons was very positive. They found it had potential to help them navigate more easily and especially for more complex cases. They also envisioned it helping with planning and allow them to more easily explain these plans to residents in the room.

7.1.4 Intraoperative Augmented Reality Video-Based Hemodynamic Analysis

We have also investigated if mobile devices are suitable for intraoperative video-based hemodynamic guidance. An existing video hemodynamic analysis software was integrated into MARIN and a proof-of-concept processing pipeline was constructed and tested. The prototype enables a user to record a video sequence, select a region of interest for analysis directly onto the AR view, on the mobile device, and display the resulting hemodynamic information in the AR view. The system was found to be suitable: videos captured were of sufficient resolution and frame rate for the analysis to work and the system is simple and can easily be integrated to the surgical workflow. Akin to the work presented in Chapter 4, this method gives control back to

the practicing surgeon and alleviates the need for external intervention, which again has the potential to streamline procedures and therefore improve outcomes.

7.1.5 Low-cost Tracking for 3D Freehand Ultrasound Reconstruction

In Chapter 6, we assessed the feasibility of using low-cost tracking hardware for 3D freehand US reconstruction, a method used to correct for brain shift intraoperatively. In order to answer that question, a phantom and protocol were devised to standardize measurements and assess accuracy of the obtained reconstruction independently in all three dimensions of space. The phantom is easy to build and the protocol easy to replicate. This protocol will thus enable other teams developing either new reconstruction methods or new tracking methods to test it in a comparable and robust setup.

It was found that optical trackers, even those that retailed at a five fold difference in price, performed similarly for this application. The same was true for both US imaging systems tested, which are retailed at a 35 fold difference in price. These findings hint that a complete hardware setup for brain shift correction can be built at a fraction of the cost of the that of the gold standard with little impact on performance.

Although, it was found that the cheapest tracking methods and hardware, *i.e.,* the image-based method and the sensor fusion method, performed rather poorly and would not be usable for this application. We do not believe this means that cheap tracking cannot be used in the surgical domain, but rather that new set-ups or algorithms that account for lower quality data need to be developed.

7.2 Future Research Directions

Many open questions remain with AR IGNS systems. The next sections details some of the most important avenues that would warrant being explored in future work.

7.2.1 Visualization

The first category of open problems is related to perception of displayed information and visualization. Relevant questions still to investigate are for example: what is the best means by which the augmented image should be displayed, and when and how should the augmentation be done. Comparative studies investigating this should be done. Further, it is well known that AR can have some risks of overwhelming the visual system of the user and distracting them from real-world if too much information is presented [53]. There is also the risk of clouding the view, or hiding a real object with a virtual one and thus even harming the surgeon's ability to perform a task [60].

There is also the question of appropriately presenting the depth of virtual objects to the user, which as underlined by Kersten-Oertel *et al.* [61], is of crucial importance, especially for vascular procedures. Additional studies investigating these points are still lacking. Finally, a question that is closely related to perception as well is that of evaluating the effect of the learning curve associated with getting acquainted with AR visualizations. Studies that have been done so far, and including those presented in this thesis, present two systems to subjects and measure their performance with both. Although, in all studies involving experienced neurosurgeons, the AR system comes at a disadvantage in the comparison. Indeed, even though the traditional guidance is perhaps less intuitive to grasp and visualize than AR can potentially be, senior practitioners have had a lot more experience with it than they have the AR counterpart. This can potentially bias results. Studies focusing specifically

on assessing how usage performance can increase with more practice are thus an interesting avenue of future work.

7.2.2 Uncertainty

Another limitation of current systems is the lack of any uncertainty estimate of the guidance information. Current state-of-the-art AR (and non-AR) systems display the virtual structures (pointer, anatomy, etc.) overlaid on the real view (or on the images in the case of traditional navigation) at the position where they are believed to be, but without specifying any error margins. This is problematic because many potential sources of error affect the accuracy of the guidance or overlay. The augmented structures are thus never exactly where they are expected to be, although this error can vary significantly. Knowledge of the uncertainty estimate could, in many situations, completely change the surgeon's usage of the system. At the moment, surgeons do not know what kind of degradation the guidance accuracy has suffered. They thus often stop using the system altogether at later stages of surgery to avoid relying on erroneous guidance, even though it may sometimes be still accurate. Conversely, they may sometimes use the system not knowing that the accuracy has degraded too much to be relied on. This later case poses an important risk on patient safety. Finding ways to appreciate in real time the uncertainty of the displayed information and communicating it effectively to the operating surgeon requires further exploration.

7.2.3 System Validation

AR technologies are still early in their development and long term evaluations of patient outcomes associated with their use compared to traditional navigation systems are needed. Thorough clinical validation is currently lacking and will be an important

research question to focus on going forward if we want these new tools to enter routine clinical practice.

Furthermore, questions remain, not only in clinical validation, but in technical validation as well. New metrics to assess system accuracy, performance and robustness need to be developed. As we can see in the table shown in Figure 10, error assessment varies a lot from one research prototype to the other. Therefore, building reproducible and accurate validation procedures to enable comparison of systems is necessary to better understand which methods are worth exploring further. This factor is not intrinsic to the systems themselves, but rather to their design and test phases. In addition to being hard to compare, most of the currently used error estimates are not truly representative of the real accuracy of developed systems. In-plane measurements do not necessarily give a meaningful 3D error estimate, since the same in-plane distance can have significantly different corresponding 3D distances depending on the distance between the camera and the scene. Additionally, in-plane measurements only assess discrepancy in that plane, but give no information about error in depth. They thus tend to underestimate the effective level of error associated with a system, which may lead to overconfidence in performance evaluation.

7.2.4 Mobile AR

This thesis explored some of the benefits that mobile device AR guidance could bring to neurosurgical interventions, but other potential benefits of mobile devices have not been explored. Positioning sensors, like the accelerometer and gyroscopes, that these devices are equipped with could perhaps be used in sensor-fusion methods to either refine existing tracking methods or even create new ones alleviating the need for an external device altogether. Recently, more and more mobile devices are equipped with depth sensors which could also be used, either for standalone tracking or in

combination with other sources of information. The results of Chapter 6 seem to indicate that methods relying on those sensors still require improvements to meet the requirements of neurosurgery. Although, depth sensors only recently reached affordability and a sufficiently high resolution to envision using them in surgery, development of these methods is thus still in its infancy. We believe that even though our results were rather disappointing, there is much potential for improvement.

7.3 Outlook

In this dissertation we have shown that novel low-cost devices, in particular mobile devices, have the potential to positively impact neurosurgical guidance systems through more intuitive and simpler visualization and interaction. As well, we have shown that costs can be drastically cut in certain places when choosing system hardware with very little impact on system accuracy. This thesis thus paves the way to building a complete low-cost neuronavigation system. Such a system would have a high impact on global health as it would democratize access to higher quality neurosurgical care for millions of patients.

Bibliography

[1] J. A. Aarli, T. Dua, A. Janca, and A. Muscetta. Neurological Disorders: Public Health Challenges. Technical report, World Health Organization, 2006.

[2] M. Asselin, A. Lasso, T. Ungi, and G. Fichtinger. Towards webcam-based tracking for interventional navigation. In B. Fei and R. J. W. III, editors, *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10576, pages 534 – 543. International Society for Optics and Photonics, SPIE, 2018.

[3] E.-O. Backlund. A new instrument for stereotaxic brain tumour biopsy. technical note. *Acta chirurgica Scandinavica*, 137(8):825–827, 1971.

[4] E.-O. Backlund. Studies on craniopharyngiomas. 3. stereotaxic treatment with intracystic yttrium-90. *Acta chirurgica Scandinavica*, 139(3):237, 1973.

[5] M. Bergstrom and T. Greitz. Stereotaxic computed tomography. *American Journal of Roentgenology*, 127(1):167–170, 1976.

[6] L. Besharati Tabrizi and M. Mahvash. Augmented reality–guided neurosurgery: accuracy and intraoperative application of an image projection technique. *Journal of neurosurgery*, 123(1):206–211, 2015.

[7] P. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[8] R. Bieck, S. Franke, T. Neumuth, and D. Lindner. Computer-assisted neurosurgery: an interaction concept for a tablet-based surgical assistance system. In *14th Annual Conf. the German Society of Computer and Robotic Assisted Surgery, Bremen, Germany*, 2015.

[9] Brainlab. Neurosurgery products. https://www.brainlab.com/en/surgery-products/overview-neurosurgery-products. Accessed: 2021-06-21.

[10] J. Brooke. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[11] I. Cabrilo, P. Bijlenga, and K. Schaller. Augmented reality in the surgery of cerebral aneurysms: a technical report. *Operative Neurosurgery*, 10(2):252–261, 2014.

[12] A. Cao, R. C. Thompson, P. Dumpuri, B. M. Dawant, R. L. Galloway, S. Ding, and M. I. Miga. Laser range scanning for image-guided neurosurgery: investigation of image-to-physical space registrations. *Medical physics*, 35(4):1593–1605, 2008.

[13] G. Carbajal, A. Lasso, Á. Gómez, and G. Fichtinger. Improving N-wire phantom-based freehand ultrasound calibration. *International Journal of Computer Assisted Radiology and Surgery*, 8(6):1063–1072, 2013.

[14] M. Carbone, R. Piazza, and S. Condino. Commercially Available Head-Mounted Displays Are Unsuitable for Augmented Reality Surgical Guidance : A Call for Focused Research for Surgical Applications. *Surgical Innovation*, 27(3):254–255, 2020.

[15] F. Cenni, D. Monari, K. Desloovere, E. Aertbeliën, S. H. Schless, and H. Bruyninckx. The reliability and validity of a clinical 3D freehand ultrasound system. *Computer Methods and Programs in Biomedicine*, 136:179–187, 2016.

[16] Y.-Z. Chang, J.-F. Hou, Y. H. Tsao, and S.-T. Lee. Application of real-time single camera SLAM technology for image-guided targeting in neurosurgery. *Applications of Digital Image Processing XXXV*, 8499(October 2012):84992D, 2012.

[17] O. Clatz, H. Delingette, I. F. Talos, A. J. Golby, R. Kikinis, F. A. Jolesz, N. Ayache, and S. K. Warfield. Robust nonrigid registration to capture brain shift from intraoperative MRI. *IEEE Transactions on Medical Imaging*, 24(11):1417–1427, 2005.

[18] S. Condino, M. Carbone, R. Piazza, M. Ferrari, and V. Ferrari. Perceptual Limits of Optical See-Through Visors for Augmented Reality Guidance of Manual Tasks. *IEEE Transactions on Biomedical Engineering*, 67(2):411–419, 2019.

[19] W. Deng, F. Li, M. Wang, and Z. Song. Easy-to-use augmented reality neuronavigation using a wireless tablet PC. *Stereotactic and functional neurosurgery*, 92(1):17–24, 2014.

[20] M. C. Dewan, A. Rattani, G. Fieggen, M. A. Arraez, F. Servadei, F. A. Boop, W. D. Johnson, B. C. Warf, and K. B. Park. Global neurosurgery: The current capacity and deficit in the provision of essential neurosurgical care. Executive summary of the global neurosurgery initiative at the program in global surgery and social change. *Journal of Neurosurgery*, 130(4):1055–1064, 2019.

[21] M. C. Dewan, A. Rattani, S. Gupta, R. E. Baticulon, Y. C. Hung, M. Punchak, A. Agrawal, A. O. Adeleye, M. G. Shrime, A. M. Rubiano, J. V. Rosenfeld, and K. B. Park. Estimating the global incidence of traumatic brain injury. *Journal of Neurosurgery*, 130(4):1080–1097, 2019.

[22] N. L. Dorward, O. Alberti, B. Velani, F. A. Gerritsen, W. F. J. Harkness, N. D. Kitchen, and D. G. T. Thomas. Postimaging brain distortion: magnitude, correlates, and impact on neuronavigation. *Journal of Neurosurgery*, 88(4):656–662, 1998.

[23] S. Drouin, M. Kersten-Oertel, S. J. S. Chen, and D. L. Collins. A realistic test and development environment for mixed reality in neurosurgery. *Lecture Notes in Computer Science*, 7264 LNCS:13–23, 2012.

[24] S. Drouin, M. Kersten-Oertel, and D. L. Collins. Interaction-based registration correction for improved augmented reality overlay in neurosurgery. In *Workshop on Augmented Environments for Computer-Assisted Interventions*, pages 21–29. Springer, 2015.

[25] S. Drouin, A. Kochanowska, M. Kersten-Oertel, I. J. Gerard, R. Zelmann, D. De Nigris, S. Bériault, T. Arbel, D. Sirhan, A. F. Sadikot, J. A. Hall, D. S. Sinclair, K. Petrecca, R. F. DelMaestro, and D. L. Collins. IBIS: an OR ready open-source platform for image-guided neurosurgery. *International Journal of Computer Assisted Radiology and Surgery*, 12(3):363–378, 2016.

[26] P. J. Edwards, D. J. Hawkes, D. L. Hill, D. Jewell, R. Spink, A. Strong, and M. Gleeson. Augmentation of reality using an operating microscope for otolaryngology and neurosurgical guidance. *Computer Aided Surgery*, 1(3):172–178, 1995.

[27] P. J. Edwards, A. P. King, D. J. Hawkes, O. Fleig, C. R. Maurer, D. L. Hill, M. R. Fenlon, D. A. De Cunha, R. P. Gaston, S. Chandra, J. Mannss, A. J. Strong, M. J. Gleeson, and T. C. Cox. Stereo augmented reality in the surgical microscope. *Studies in Health Technology and Informatics*, 62:102–108, 1999.

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

[29] B. Eftekhar. A Smartphone App to Assist Scalp Localization of Superficial Supratentorial Lesions - Technical Note. *World Neurosurgery*, 85:359–363, 2016.

[30] G. Eggers, J. Mühling, and R. Marmulla. Image-to-patient registration techniques in head surgery. *International Journal of Oral and Maxillofacial Surgery*, 35(12):1081–1095, 2006.

[31] F. Faul, E. Erdfelder, A. Buchner, and A. G. Lang. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4):1149–1160, 2009.

[32] E. M. Friets, J. W. Strohbehn, J. F. Hatch, and D. W. Roberts. A frameless stereotaxic operating microscope for neurosurgery. *IEEE Transactions on Biomedical Engineering*, 36(6):608–617, 1989.

[33] B. Fuerst, W. Wein, M. Müller, and N. Navab. Automatic ultrasound-MRI registration for neurosurgery using the 2D and 3D LC2Metric. *Medical Image Analysis*, 18(8):1312–1319, 2014.

[34] R. L. Galloway Jr. The process and development of image-guided procedures. *Annual Review of Biomedical Engineering*, 3(1):83–108, 2001.

[35] R. L. Galloway Jr. and T. M. Peters. Overview and history of image-guided interventions. In *Image-Guided Interventions*, pages 1–21. Springer, 2008.

[36] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014.

[37] I. J. Gerard and D. L. Collins. An analysis of tracking error in image-guided neurosurgery. *International Journal of Computer Assisted Radiology and Surgery*, 10(10):1579–1588, 2015.

[38] I. J. Gerard, M. Kersten-Oertel, S. Drouin, J. A. Hall, K. Petrecca, D. De Nigris, D. A. Di Giovanni, T. Arbel, and D. L. Collins. Combining intraoperative ultrasound brain shift correction and augmented reality visualizations: a pilot study of eight cases. *Journal of Medical Imaging*, 5(2), 2018.

[39] I. J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J. A. Hall, and D. L. Collins. Brain shift in neuronavigation of brain tumors: A review. *Medical image analysis*, 35:403–420, 2017.

[40] V. M. Gerganov, A. Samii, M. Giordano, M. Samii, and R. Fahlbusch. Two-dimensional high-end ultrasound imaging compared to intraoperative MRI during resection of low-grade gliomas. *Journal of Clinical Neuroscience*, 18(5):669–673, 2011.

[41] P. L. Gildenberg, R. Ledoux, E. Cosman, and J. Labuz. The exoscope-a frame-based video/graphics system for intraoperative guidance of surgical resection. *Stereotactic and functional neurosurgery*, 63(1-4):23–25, 1994.

[42] P. L. Gleason, R. Kikinis, D. Altobelli, W. Wells, E. Alexander III, P. M. Black, and F. A. Jolesz. Video registration virtual reality for nonlinkage stereotactic surgery. *Stereotactic and functional neurosurgery*, 63(1-4):139–143, 1994.

[43] K. H. Goodell, C. G. Cao, and S. D. Schwaitzberg. Effects of cognitive distraction on performance of laparoscopic surgical tasks. *Journal of Laparoendoscopic & Advanced Surgical Techniques*, 16(2):94–98, 2006.

[44] J. Graydon and M. W. Eysenck. Distraction and cognitive performance. *European Journal of Cognitive Psychology*, 1(2):161–179, 1989.

[45] W. E. L. Grimson, T. Lozano-Perez, W. M. Wells, G. J. Ettinger, S. J. White, and R. Kikinis. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. *IEEE Transactions on Medical Imaging*, 15(2):129–140, 1996.

[46] H. E. Gueziri, O. Rabau, C. Santaguida, and D. L. Collins. Evaluation of an Ultrasound-Based Navigation System for Spine Neurosurgery: A Porcine Cadaver Study. *Frontiers in Oncology*, 11:228, 2021.

[47] D. Guha, N. M. Alotaibi, N. Nguyen, S. Gupta, C. McFaul, and V. X. D. Yang. Augmented Reality in Neurosurgery: A Review of Current Concepts and Emerging Applications. *Canadian Journal of Neurological Sciences*, 44(3):235–245, 2017.

[48] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[49] R. L. Holloway. Registration Error Analysis for Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4):413–432, 1997.

[50] V. Horsley and R. Clarke. The structure and functions of the cebellum examined by a new method. *Brain : a journal of neurology*, 31(December):45–124, 1908.

[51] Y. Hou, L. Ma, R. Zhu, X. Chen, and J. Zhang. A Low-Cost iPhone-Assisted Augmented Reality Solution for the Localization of Intracranial Lesions. *PLOS ONE*, 11(7):1–18, 2016.

[52] G. N. Hounsfield. Computerized transverse axial scanning (tomography): Part 1. description of system. *The British journal of radiology*, 46(552):1016–1022, 1973.

[53] A. Hughes-Hallett, E. K. Mayer, H. J. Marcus, P. Pratt, S. Mason, A. W. Darzi, and J. A. Vale. Inattention blindness in surgery. *Surgical Endoscopy and Other Interventional Techniques*, 29(11):3184–3189, 2015.

[54] International Agency for Research on Cancer, World Health Organization. Global cancer observatory. `https://gco.iarc.fr/tomorrow/en/dataviz/isotype`. Accessed: 2021-06-21.

[55] S. Ji, Z. Wu, A. Hartov, D. W. Roberts, and K. D. Paulsen. Mutual-information-based image to patient re-registration using intraoperative ultrasound in image-guided neurosurgery. *Medical Physics*, 35(10):4612–4624, 2008.

[56] F. Jolesz. *Intraoperative imaging and image-guided therapy*. Springer Science & Business Media, 2014.

[57] S. R. Kantelhardt, A. Gutenberg, A. Neulen, N. Keric, M. Renovanz, and A. Giese. Video-assisted navigation for adjustment of image-guidance accuracy to slight brain shift. *Operative Neurosurgery*, 11(4):504–511, 2015.

[58] T. Kawamata, H. Iseki, T. Shibasaki, T. Hori, C. B. Heilman, G. H. Barnett, and M. L. Levy. Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors: Technical note. *Neurosurgery*, 50(6):1393–1397, 2002.

[59] P. J. Kelly, B. A. Kall, S. Goerss, and F. Earnest IV. Computer-assisted stereotaxic laser resection of intra-axial brain neoplasms. *Journal of neurosurgery*, 64(3):427–439, 1986.

[60] M. Kersten-Oertel, S. J.-S. Chen, S. Drouin, D. S. Sinclair, and D. L. Collins. Augmented reality visualization for guidance in neurovascular surgery. In *MMVR*, pages 225–229, 2012.

[61] M. Kersten-Oertel, I. J. Gerard, S. Drouin, K. Mok, D. Sirhan, D. S. Sinclair, and D. L. Collins. Augmented reality in neurovascular surgery: feasibility and first uses in the operating room. *International Journal of Computer Assisted Radiology and Surgery*, 10(11):1823–1836, 2015.

[62] M. Kersten-Oertel, I. J. Gerard, S. Drouin, K. Petrecca, J. A. Hall, and D. L. Collins. Towards augmented reality guided craniotomy planning in tumour resections. In *International Conference on Medical Imaging and Virtual Reality*, pages 163–174. Springer, 2016.

[63] M. Kersten-Oertel, P. Jannin, and D. L. Collins. The state of the art of visualization in mixed reality image guided surgery. *Computerized Medical Imaging and Graphics*, 37(2):98–112, 2013.

[64] R. Kikinis, P. Gleason, W. E. Lorensen, W. M. Wells III, W. E. L. Grimson, T. Lozano-Perez, G. J. Ettinger, S. F. White, and F. A. Jolesz. Image guidance techniques for neurosurgery. In *Visualization in Biomedical Computing 1994*, pages 537–540. International Society for Optics and Photonics, 1994.

[65] R. Kikinis, P. L. Gleason, T. M. Moriarty, M. R. Moore, E. Alexander, P. E. Stieg, M. Matsumae, W. E. Lorensen, H. E. Cline, P. M. Black, and F. A. Jolesz. Computer-assisted interactive three-dimensional planning neurosurgical procedures. *Neurosurgery*, 38(4):640–651, 1996.

[66] Y. Kosugi, E. Watanabe, J. Goto, T. Watanabe, S. Yoshimoto, K. Takakura, and J. Ikebe. An articulated neurosurgical navigation system using MRI and CT images. *IEEE Transactions on Biomedical Engineering*, 35(2):147–152, 1988.

[67] A. N. Kumar, M. I. Miga, T. S. Pheiffer, L. B. Chambless, R. C. Thompson, and B. M. Dawant. Automatic tracking of intraoperative brain surface displacements in brain tumor surgery. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 1509–1512. IEEE, 2014.

[68] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger. PLUS: Open-source toolkit for ultrasound-guided intervention systems. *IEEE Transactions on Biomedical Engineering*, 61(10):2527–2537, 2014.

[69] A. Leibinger, A. E. Forte, Z. Tan, M. J. Oldfield, F. Beyrau, D. Dini, and F. Rodriguez y Baena. Soft Tissue Phantoms for Realistic Needle Insertion: A Comparative Study. *Annals of Biomedical Engineering*, 44(8):2442–2452, 2016.

[70] L. Leksell. A stereotaxic apparatus for intracerebral surgery. *Acta Chirurgica Scandinavica*, 99(3):229–233, 1950.

[71] S. Lemes. Comparison of similar injection moulded parts by a coordinate measuring machine. *SN Applied Sciences*, 1(2):1–8, 2019.

[72] G. M. Malham and T. Wells-Quinn. What should my hospital buy next?—Guidelines for the acquisition and application of imaging, navigation, and robotics for spine surgery. *Journal of Spine Surgery*, 5(1):155–165, 2019.

[73] H. J. Marcus, P. Pratt, A. Hughes-Hallett, T. P. Cundy, A. P. Marcus, G. Z. Yang, A. Darzi, and D. Nandi. Comparative effectiveness and safety of image guidance systems in neurosurgery: A preclinical randomized study. *Journal of Neurosurgery*, 123(2):307–313, 2015.

[74] Y. Masutani, T. Dohi, F. Yamane, H. Iseki, and K. Takakura. Augmented reality visualization system for intravascular neurosurgery. *Computer Aided Surgery*, 3(5):239–247, 1998.

[75] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 2003.

[76] Medtronic plc. Surgical navigation systems. `http://www.medtronic.com/us-en/healthcare-professionals/products/neurological/surgical-navigation-systems.html`. Accessed: 2021-06-21.

[77] P. Menei, J.-P. Benoit, M. Boisdron-Celle, D. Fournier, P. Mercier, and G. Guy. Drug targeting into the central nervous system by stereotactic implantation of biodegradable microspheres. *Neurosurgery*, 34(6):1058–1064, 1994.

[78] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari, and V. Ferrari. Augmented reality in neurosurgery: a systematic review. *Neurosurgical review*, 40(4):537–548, 2017.

[79] L. Mercier, R. F. Del Maestro, K. Petrecca, A. Kochanowska, S. Drouin, C. X. B. Yan, A. L. Janke, S. J.-S. Chen, and D. L. Collins. New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: system description and validation. *International Journal of Computer Assisted Radiology and Surgery*, 6(4):507–522, 2011.

[80] L. Mercier, T. Langø, F. Lindseth, and D. L. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine and Biology*, 31(4):449–471, 2005.

[81] M. I. Miga. Computational Modeling for Enhancing Soft Tissue Image Guided Surgery: An Application in Neurosurgery. *Annals of Biomedical Engineering*, 44(1):128–138, 2016.

[82] M. I. Miga, K. Sun, I. Chen, L. W. Clements, T. S. Pheiffer, A. L. Simpson, and R. C. Thompson. Clinical evaluation of a model-updated image-guidance approach to brain shift compensation: experience in 16 cases. *International Journal of Computer Assisted Radiology and Surgery*, 11(8):1467–1474, 2015.

[83] M. I. Miga, K. Sun, I. Chen, L. W. Clements, T. S. Pheiffer, A. L. Simpson, and R. C. Thompson. Clinical evaluation of a model-updated image-guidance approach to brain shift compensation: experience in 16 cases. *International journal of computer assisted radiology and surgery*, 11(8):1467–1474, 2016.

[84] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–293. International Society for Optics and Photonics, 1995.

[85] R. C. Miner. Image-Guided Neurosurgery. *Journal of Medical Imaging and Radiation Sciences*, 48(4):328–335, 2017.

[86] M. H. Mobasheri, M. Johnston, U. M. Syed, D. King, and A. Darzi. The uses of smartphones and tablet devices in surgery: A systematic review of the literature. *Surgery*, 158(5):1352–1371, 2015.

[87] A. Moiyadi and P. Shetty. Objective assessment of utility of intraoperative ultrasound in resection of central nervous system tumors: A cost-effective tool for intraoperative navigation in neurosurgery. *Journal of neurosciences in rural practice*, 2(1):4, 2011.

[88] F. Morin, H. Courtecuisse, I. Reinertsen, F. Le Lann, O. Palombi, Y. Payan, and M. Chabanas. Brain-shift compensation using intraoperative ultrasound and constraint-based biomechanical simulation. *Medical Image Analysis*, 40:133–153, 2017.

[89] M. Müller, M.-C. Rassweiler, J. Klein, A. Seitel, M. Gondan, M. Baumhauer, D. Teber, J. J. Rassweiler, H.-P. Meinzer, and L. Maier-Hein. Mobile augmented reality for computer-assisted percutaneous nephrolithotomy. *International journal of computer assisted radiology and surgery*, 8(4):663–675, 2013.

[90] F. Mundinger and T. Hoefer. Protracted long-term irradiation of inoperable midbrain tumours by stereotactic curie-therapy using iridium-192. In *Advances in Stereotactic and Functional Neurosurgery*, pages 93–100. Springer, 1974.

[91] A. Nabavi, P. M. Black, D. T. Gering, C.-F. Westin, V. Mehta, R. S. Pergolizzi, M. Ferrant, S. K. Warfield, N. Hata, R. B. Schwartz, W. M. Wells III, R. Kikinis, and F. A. Jolesz. Serial Intraoperative MR Imaging of Brain Shift. *Neurosurgery*, 48(4):787–798, 2001.

[92] D. D. Nigris, D. L. Collins, and T. Arbel. Fast rigid registration of pre-operative magnetic resonance images to intra-operative ultrasound for neurosurgery based on high confidence gradient orientations. *International Journal of Computer Assisted Radiology and Surgery*, 8(4):649–661, 2013.

[93] C. Nimsky, O. Ganslandt, P. Hastreiter, and R. Fahlbusch. Intraoperative compensation for brain shift. *Surgical Neurology*, 56(6):357–364, 2001.

[94] S. Ohue, Y. Kumon, S. Nagato, S. Kohno, H. Harada, K. Nakagawa, K. Kikuchi, H. Miki, and T. Ohnishi. Evaluation of Intraoperative Brain Shift Using an Ultrasound-Linked Navigation System for Brain Tumor Surgery. *Neurologia medico-chirurgica*, 50(4):291–300, 2010.

[95] D. G. Pearson and A. Sahraie. Oculomotor control and the maintenance of spatially and temporally distributed events in visuo-spatial working memory. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 56 A(7):1089–1111, 2003.

[96] V. M. Pereira, I. Smit-Ockeloen, O. Brina, D. Babic, M. Breeuwer, K. Schaller, K. O. Lovblad, and D. Ruijters. Volumetric measurements of brain shift using intraoperative cone-beam computed tomography: Preliminary study. *Operative Neurosurgery*, 12(4):4–12, 2016.

[97] S. Pieper, M. Halle, and R. Kikinis. 3D Slicer. *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano*, 1(March 2015):632–635, 2004.

[98] B. R. Postle, C. Idzikowski, S. D. Sala, R. H. Logie, and A. D. Baddeley. The selective disruption of spatial working memory by eye movements. *The quarterly Journal of experimental Psychology*, 59(1):100–120, 2006.

[99] I. Reinertsen and M. Descoteaux. Validation of vessel-based registration for correction of brain shift. *Direct*, 11(4):374–388, 2007.

[100] I. Reinertsen, F. Lindseth, C. Askeland, D. H. Iversen, and G. Unsgård. Intraoperative correction of brain-shift. *Acta Neurochirurgica*, 156(7):1301–1310, 2014.

[101] I. Reinertsen, F. Lindseth, C. Askeland, D. H. Iversen, and G. Unsgård. Intraoperative correction of brain-shift. *Acta Neurochirurgica*, 156(7):1301–1310, 2014.

[102] I. Reinertsen, F. Lindseth, G. Unsgaard, and D. L. Collins. Clinical validation of vessel-based registration for correction of brain-shift. *Medical Image Analysis*, 11(6):673–684, 2007.

[103] T. Riechert and M. Wolff. Über ein neues zielgerät zur intrakraniellen elektrischen ableitung und ausschaltung. *European Archives of Psychiatry and Clinical Neuroscience*, 186(2):225–230, 1951.

[104] M. Riva, P. Hiepe, M. Frommert, I. Divenuto, L. G. Gay, T. Sciortino, M. C. Nibali, M. Rossi, F. Pessina, and L. Bello. Intraoperative computed tomography and finite element modelling for multimodal image fusion in brain surgery. *Operative Neurosurgery*, 18(5):531–541, 2019.

[105] H. Rivaz and D. L. Collins. Deformable registration of preoperative MR, pre-resection ultrasound, and post-resection ultrasound images of neurosurgery. *International Journal of Computer Assisted Radiology and Surgery*, 2015.

[106] D. W. Roberts, J. W. Strohbehn, J. F. Hatch, W. Murray, and H. Kettenberger. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *Journal of neurosurgery*, 65(4):545–549, 1986.

[107] A. L. Roethe, J. Rösler, P. Vajkoczy, and T. Picht. Current experiences and future perspectives for augmented reality visualization in navigated

neurosurgical interventions. In *Computer Assisted Radiology and Surgery*, page S79, Berlin, Germany, 2018.

[108] R. Rohling, A. Gee, and L. Berman. A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical Image Analysis*, 3(4):339–359, 1999.

[109] W. C. Röntgen. On a new kind of rays. *Science*, 3(59):227–231, 1896.

[110] D. Ruijters, B. M. ter Haar Romeny, and P. Suetens. GPU-accelerated elastic 3D image registration for intra-surgical applications. *Computer Methods and Programs in Biomedicine*, 103(2):104–112, 2011.

[111] S. Schnaudigel, C. Preul, T. Ugur, H. J. Mentzel, O. W. Witte, M. Tittgemeyer, and G. Hagemann. Positional brain deformation visualized with magnetic resonance morphometry. *Neurosurgery*, 66(2):376–384, 2010.

[112] C. Schulz, S. Waldeck, and U. M. Mauer. Intraoperative image guidance in neurosurgery: development, current indications, and future trends. *Radiology research and practice*, vol. 2012, 2012.

[113] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced Medical Displays: A Literature Review of Augmented Reality. *Journal of Display Technology*, 4(4):451–467, 2008.

[114] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. Nagelhus Hernes. Freehand 3D Ultrasound Reconstruction Algorithms-A Review. *Ultrasound in Medicine and Biology*, 33(7):991–1009, 2007.

[115] E. A. Spiegel, H. T. Wycis, M. Marks, and A. J. Lee. Stereotaxic apparatus for operations on the human brain. *Science*, 106(2754):349–350, 1947.

[116] L. H. Stieglitz, J. Fichtner, R. Andres, P. Schucht, A. K. Krähenbühl, A. Raabe, and J. Beck. The silent loss of neuronavigation accuracy: A systematic retrospective analysis of factors influencing the mismatch of frameless stereotactic systems in cranial neurosurgery. *Neurosurgery*, 72(5):796–807, 2013.

[117] H. Sun, D. W. Roberts, H. Farid, Z. Wu, A. Hartov, and K. D. Paulsen. Cortical surface tracking using a stereoscopic operating microscope. *Operative Neurosurgery*, 56(suppl1):86–97, 2005.

[118] R. Tagaytayan, A. Kelemen, and C. Sik-Lanyi. Augmented reality in neurosurgery. *Archives of Medical Science*, 12(1), 2016.

[119] J. Talairach, H. Hecaen, M. David, M. Monnier, and J. Deajuriaguerra. Recherches sur la coagulation thérapeutique des structures sous-corticales chez l'homme. *Revue neurologique*, 81(1):4–24, 1949.

[120] J. Tokuda, G. S. Fischer, X. Papademetris, Z. Yaniv, L. Ibanez, P. Cheng, H. Liu, J. Blevins, J. Arata, A. J. Golby, T. Kapur, S. Pieper, E. C. Burdette, G. Fichtinger, C. M. Tempany, and N. Hata. OpenIGTLink: An open network protocol for image-guided therapy environment. *International Journal of Medical Robotics and Computer Assisted Surgery*, 5(4):423–434, 2009.

[121] T. Ungi, A. Lasso, and G. Fichtinger. Open-source platforms for navigated image-guided interventions. *Medical Image Analysis*, 33:181–186, 2016.

[122] G. Unsgaard, S. Ommedal, T. Muller, A. Gronningsaeter, and T. A. N. Hernes. Neuronavigation by intraoperative 3D ultrasound, Initial experiences during brain tumor resections. *Neurosurgery*, 50(4):804–812, 2002.

[123] G. Unsgaard, O. M. Rygh, T. Selbekk, T. B. Müller, F. Kolstad, F. Lindseth, and T. A. N. Hernes. Intra-operative 3D ultrasound in neurosurgery. *Acta Neurochirurgica*, 148(3):235–253, 2006.

[124] R. Vassallo, A. Rankin, S. P. Lownie, H. Fukuda, H. Kasuya, B. W. Lo, T. Peters, and Y. Xiao. Determining blood flow direction from short neurovascular surgical microscope videos. *Healthcare Technology Letters*, 6(6):191–196, 2019.

[125] J. P. Wachs. Gaze, posture and gesture recognition to minimize focus shifts for intelligent operating rooms in a collaborative support system. *International Journal of Computers Communications & Control*, 5(1):106–124, 2010.

[126] M. N. Wang and Z. J. Song. Classification and analysis of the errors in neuronavigation. *Neurosurgery*, 68(4):1131–1143, 2011.

[127] F. A. Wanis, L. Wessels, M. H. Reinges, E. Uhl, and A. Jödicke. Technical accuracy of the integration of an external ultrasonography system into a navigation platform: effects of ultrasonography probe registration and target detection. *Acta Neurochirurgica*, 160(2):305–316, 2018.

[128] E. Watanabe, M. Satoh, T. Konno, M. Hirai, and T. Yamaguchi. The Trans-Visible Navigator: A See-Through Neuronavigation System Using Augmented Reality. *World Neurosurgery*, 87:399–405, 2016.

[129] E. Watanabe, T. Watanabe, S. Manaka, Y. Mayanagi, and K. Takakura. Three-dimensional digitizer (neuronavigator): New equipment for computed tomography-guided stereotaxic surgery. *Surgical Neurology*, 27(6):543–547, 1987.

[130] V. Weerdesteyn, A. Schillings, G. Van Galen, and J. Duysens. Distraction affects the performance of obstacle avoidance during walking. *Journal of motor behavior*, 35(1):53–63, 2003.

[131] W. S. R. Wirtz, F. Albert, M. Schwaderer, C. Heuer, A. Staubert, V. Tronnier, M. Knauth, and S. Kunze. The benefit of neuronavigation for neurosurgery analyzed by its impact on glioblastoma surgery. *Neurological Research*, 22(4):354–360, 2000.

[132] A. Wittek, K. Miller, R. Kikinis, and S. K. Warfield. Patient-specific model of brain deformation: Application to medical image registration. *Journal of Biomechanics*, 40(4):919–929, 2007.

[133] Y. Xiao, J. C. Lau, D. Hemachandra, G. Gilmore, A. R. Khan, and T. M. Peters. Image Guidance in Deep Brain Stimulation Surgery to Treat Parkinson's Disease: A Comprehensive Review. *IEEE Transactions on Biomedical Engineering*, 68(3):1024–1033, 2021.

[134] Y. Xiao, H. Rivaz, H. Kasuya, S. Yokosako, C. Mindru, J. Teitelbaum, D. Sirhan, D. Sinclair, M. Angle, and B. W. Lo. Intra-operative video characterization of carotid artery pulsation patterns in case series with post-endarterectomy hypertension and hyperperfusion syndrome. *Translational stroke research*, 9(5):452–458, 2018.

[135] Y. Xiao, C. X. B. Yan, S. Drouin, D. De Nigris, A. Kochanowska, and D. L. Collins. User-friendly freehand ultrasound calibration using Lego bricks and automatic registration. *International Journal of Computer Assisted Radiology and Surgery*, 11(9):1703–1711, 2016.

[136] X. Yang, L. W. Clements, S. Narasimhan, M. Luo, R. C. Thompson, B. M. Dawant, and M. I. Miga. Stereovision-based integrated system for point cloud reconstruction and simulated brain shift validation. *Journal of Medical Imaging*, 4:4–9, 2017.

[137] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

[138] Y. Zhou, G. Lelu, B. Labbé, G. Pasquier, P. Le Gargasson, A. Murienne, and L. Launay. A Novel 2D Augmented Reality Ultrasound Framework using an RGB-D Camera and a 3D-printed Marker. *International Journal of Computer Assisted Radiology and Surgery*, 16(suppl1):118 – 119, 2021.

APPENDIX A

User Study Questionnaires

A.1 Focus Shift Study Questionnaire

Navigation Systems Testing

*Required

1. Age *

2. Sex *

Mark only one oval.

- M
- F
- Other
- Prefer not to say

3. Occupation/Field of Study *

4. Field of work or study *

5. Do you know the basic functionality of an IGS system? *

Mark only one oval.

- Yes
- No

6. Have you every used/played around with an IGS system before? *

Mark only one oval.

- ◯ Yes
- ◯ No

7. How would you describe your prior experience with regards to usage of IGS navigation systems? *

Mark only one oval.

	1	2	3	4	5	
Very little	◯	◯	◯	◯	◯	A lot

Image-Guided Surgery Navigation System

This is the system with the scan presented on the computer screen.

Mental Demand

How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

8. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Physical Demand

How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

9. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Temporal Demand

How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

10. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

11. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

12. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

13. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Desktop Augmented Reality System

This is the system with the augmented live video feed presented on the computer screen.

Mental Demand

How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

14. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Physical Demand

How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

15. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Temporal Demand

How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

16. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

17. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

18. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

19. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Mobile Augmented Reality System

This is the system with the augmented live video feed presented on the phone.

Mental Demand

How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

20. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Physical Demand

How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

21. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Temporal Demand

How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

22. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

23. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

24. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

25. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Impressions

26. Did you use the pointer for Mobile AR?

Mark only one oval.

- Yes
- No
- Maybe

27. Did you use the pointer for Desktop AR?

Mark only one oval.

- Yes
- No
- Maybe

28. Did you use the pointer for traditional IGNS?

Mark only one oval.

- Yes
- No
- Maybe

29. Which of the systems did you find the most intuitive? to use? *

Mark only one oval.

- Image-Guided Surgery Navigation System
- Desktop Augmented Reality System
- Mobile Augmented Reality System
- No preference

30. Which of the systems did you feel most comfortable using? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Desktop Augmented Reality System
- ◯ Mobile Augmented Reality System
- ◯ No preference

31. Which of the systems did you feel you were being the most accurate with? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Desktop Augmented Reality System
- ◯ Mobile Augmented Reality System
- ◯ Can't say

32. Which system do you think took you the longest amount of time to draw the contour of the tumour? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Desktop Augmented Reality System
- ◯ Mobile Augmented Reality System
- ◯ Not sure.

33. Overall which of the systems did you prefer? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Desktop Augmented Reality System
- ◯ Mobile Augmented Reality System
- ◯ No preference

34. Do you have any additional comments or feedback?

This content is neither created nor endorsed by Google.

Google Forms

A.2 Manual Registration Correction Assessment Study Questionnaire

AR manual reregistration

*Required

1. SubjectID *

2. Age *

3. Sex *

   Mark only one oval.

   - ◯ M
   - ◯ F
   - ◯ Other
   - ◯ Prefer not to say
   - ◯ Other: _____

4. Occupation / Field of Study *

5. Were you familiar with Image-guidance navigation systems before this study?

   Mark only one oval.

   - ◯ Yes
   - ◯ No
   - ◯ Somewhat

6. Were you familiar with the concept of augmented reality before this study?

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Somewhat

7. Were you familiar with the concept of patient-to-image registration before this study? *

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Somewhat

Usability

8. I think that I would like to use this app frequently.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

9. I found this app unnecessarily complex.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

10. I thought this app was easy to use.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

11. I think that I would need assistance to be able to use this app.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

12. I found the various functions in this app were well integrated.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

13. I thought there was too much inconsistency in this app.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

14. I would imagine that most people would learn to use this app very quickly.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

15. I found this app very cumbersome/awkward to use.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

16. I felt very confident using this app.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

17. I needed to learn a lot of things before I could get going with this app.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

18. I feel like I was able to accurately reregister images.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

19. I could easily understand the anatomy.

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

Final comments

20. Were some features hard to use? If so, how could they have been improved?

21. Did you feel some problems in the app impacted your accuracy? If so, how could it hav been improved?

A.3 Marin Usability Assessment Study Questionnaire

MARIN Usability Questionnaire

*Required

1. Subject ID *

2. Age *

3. Sex *

Mark only one oval.

- ◯ M
- ◯ F
- ◯ Other
- ◯ Prefer not to say

4. Occupation/Field of Work or Study *

5. Did you know the basic functionality of an IGS system before? *

Mark only one oval.

- ◯ Yes
- ◯ No

6. Had you every used/played around with an IGS system before? *

Mark only one oval.

- Yes
- No

7. How would you describe your prior experience with regards to usage of IGS navigation systems? *

Mark only one oval.

	1	2	3	4	5	
Very little						A lot

8. How would you describe your prior experience with regards to usage of augmented reality? *

Mark only one oval.

	1	2	3	4	5	
Very little						A lot

9. How would you describe your prior experience with regards to reading scans or looking at 3D anatomical data? *

Mark only one oval.

	1	2	3	4	5	
Very little						A lot

Image-Guided Surgery Navigation System

This is the system with the scan slices and the 3D model.

Mental Demand

How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

10. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Physical Demand

How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

11. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Temporal Demand

How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

12. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

13. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

14. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

15. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Mobile Augmented Reality System

This is the system with the augmented live video feed presented on the phone.

Mental Demand

How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

16. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Physical Demand

How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

17. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low											Very High

Temporal Demand

How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

18. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

19. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

20. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Frustration

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

21. *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Very Low	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Very High

Impressions

22. Which of the systems did you find the most intuitive to use? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Mobile Augmented Reality System
- ◯ No preference

23. Which of the systems did you feel most comfortable using? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Mobile Augmented Reality System
- ◯ No preference

24. Which of the systems did you feel you were being the most accurate with? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Mobile Augmented Reality System
- ◯ Can't say

25. Which system do you think took you the longest amount of time to get to the target? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Mobile Augmented Reality System
- ◯ Not sure.

26. Overall which of the systems did you prefer? *

Mark only one oval.

- ◯ Image-Guided Surgery Navigation System
- ◯ Mobile Augmented Reality System
- ◯ No preference

27. Do you have any additional comments or feedback?