



university of
 groningen



umcg

Usability Evaluation of the Kinect in Aiding Surgeon-Computer Interaction

*A study on the implementation, evaluation and
improvement of gesture-based interaction in the
operating room*

Sebastiaan Michael Stuij
May 2013

Master Thesis
Human-Machine Communication
University of Groningen, The Netherlands

Internal supervisor:
Dr. F. Cnossen (Artificial Intelligence, University of Groningen)

External supervisor:
Dr. P.M.A. van Ooijen (Department of Radiology, University
Medical Centre Groningen)

Keywords

Gesture, Gesture-based interaction, Natural user interface, Usability, Surgeon, Operating Room, Medical-image viewer, Kinect

Abstract

Interest in Gesture-based interaction in the operating room (OR) environment is rising. The main advantage of introducing such an interface in the OR is that it enables direct interaction between computer and surgeon while ensuring asepsis, as opposed to asking an assistant to interact with the patient's medical images. The purpose of this study was to determine whether a modern gesture-based interface using the Kinect is feasible and desirable during surgical procedures.

After an extensive exploratory research phase including OR observations, interviews with surgeons and a questionnaire, a user-based usability evaluation was conducted with the open-source medical imaging toolkit MITO and the Microsoft Kinect. Healthcare professionals were asked to conduct prototypical tasks in a simulated OR environment in the University Medical Centre of Groningen. Obtained performance and usability measures were compared to a control condition where the participant gave instructions to an assistant, comparable to the current OR situation. Results of the usability evaluation indicated that surgeons were generally positive about gesture-based interaction and would like to use the tested system. Performance measures indicated that the current system was generally slower in executing the prototypical tasks compared to asking an assistant. However this was during their first encounter with such a novel technique; an expert user showed significant faster completion times. Another limitation of using the Kinect as gesture-based interaction technique is its reduced accuracy while conducting measurements on medical images for example.

Due to the importance of accurate selection in clinical image viewers a second study was conducted on different selection techniques in order to determine which technique is most accurate and appropriate for gesture-based selection. Two popular selection techniques: 'Dwell' and 'Push' were compared to the current mouse condition. Furthermore two different spatial resolutions were compared due to the importance of a small interaction space above the patient. Results from this experiment indicated that the tested techniques are significantly less accurate and more time-costly than the mouse control condition. However there was a significant effect between the two different spatial resolutions, indicating the importance of higher resolution depth-cameras.

Finally suggestions for usability improvements for the test-case system were proposed and important guidelines for future gesture-based interaction systems in the operating room.

From these results we can conclude that the concept of gesture-based interaction using low-cost commercially available hardware, such as the Kinect, is feasible for operating room purposes. Although the accuracy is lower and execution times are slower compared to the current situation in which the surgeon directs an assistant, surgeons rate the usability of the tested system high, and would already prefer to use this system than asking an assistant due to the direct and sterile form of interaction. Furthermore training and future technological innovations such as higher resolution depth-camera's can possibly improve the performance of gesture-based interaction.

“If I had asked people what they wanted, they would have said faster horses.”

— Henry Ford

Acknowledgements

I would like to thank the following people:

Fokke Cnossen, for her detailed feedback and good advice during the whole process ultimately leading to better academic writing and research skills.

Peter van Ooijen, for his immediate feedback during my daily activities at the UMCG and his practical and useful insights during my whole project.

Luigi Gallo and Alessio Placitelli, for their cooperation on this project and assistance on the Mito software.

Jetse Goris and Henk ten Cate Hoedemaker, for providing interesting insights and the possibility to present my project to surgeons in the UMCG.

Paul Jutte & Jasper Gerbers, for their tremendous help, insights and generosity by letting me observe their surgical procedures and getting me in contact with other surgeons.

Sip Zwerver, for letting me use the Skills Centre in the hospital as test location.

My fellow students and friends Lennart, Stephan, Jarno, Michiel, Jeffrey, Thomas and Wiard, who have provided me with countless tips and motivated me throughout the project.

Ilse, for always being there for me and always motivating me to keep going on.

And last but not least my family, who have always supported me.

Contents

Keywords	2
Abstract.....	3
Acknowledgements.....	5
Contents.....	6
List of Acronyms	9
Acronyms	9
Chapter 1.	10
Introduction	10
1.1 Problem description.....	10
1.3 This thesis	11
1.4 Research question and objectives.....	11
1.5 Thesis organization	12
Chapter 2.	13
Theoretical background.....	13
2.1 Human-computer interaction and usability.....	13
2.2 Gesture-based interaction	15
Gestures	15
Gesture interaction challenges	16
Gesture-based selection.....	16
The Kinect.....	17
2.3 Related work	18
Usability research on gesture-based interaction.....	18
Research on gesture-based selection techniques.....	19
Research on gesture-based interaction in the operating room.....	20
Chapter 3.	23
Exploratory Study.....	23
3.1 Interviews and operating room observations	23
3.2 Online questionnaire.....	25
Chapter 4.	27
Practical background	27
4.1 Current clinical medical image viewer.....	27
4.2 Gesture-based medical image viewer	28
Chapter 5.	32
Test case usability evaluation	32
5.1 Method.....	32
Participants	32
Test environment	32
Apparatus.....	33
Materials.....	33
Design and procedure.....	35
Data analysis.....	36
5.2 Performance measure results	37

Task execution time.....	38
Accuracy	40
Number of misrecognized gestures.....	41
Number of incorrectly issued gestures	42
5.3 Usability results	43
Pre-experimental familiarity with gesture-based interaction.....	43
System Usability Scale (SUS) response	44
Gesture specific questionnaire response	45
Gesture-based interaction questionnaire response	46
Open question responses.....	47
5.4 Summary.....	48
Chapter 6.	50
Performance evaluation of gesture-based selection techniques	50
6.1 Method.....	50
Participants	50
Test environment	50
Apparatus.....	50
Materials.....	51
Design and procedure.....	51
Data analysis.....	52
6.2 Performance measure results	54
Selection time.....	55
Accuracy	56
Number of errors.....	57
Throughput	58
6.3 Usability results	59
Gesture-based interaction questionnaire response	59
6.4 Summary.....	59
Chapter 7.	60
Usability recommendations.....	60
7.1 Observed test case usability issues	60
Functional improvements.....	60
Gesture improvements.....	61
Interface improvements	62
Miscellaneous improvements	62
7.2 Guidelines for gesture-based interaction systems for operating room purposes	63
Technical design considerations.....	63
Interface design considerations.....	66
Chapter 8.	67
Discussion	67
8.1 Summary of results.....	67
What do surgeons expect from gesture-based interaction?	67
Can the Kinect serve as a better way of interacting with medical images as opposed to asking an assistant?.....	68
How can gesture-based interaction in the OR be improved?	69
8.2 Implications.....	69
8.3 Limitations.....	70
8.4 Future developments	70
Chapter 9.	72

Conclusion.....	72
Bibliography.....	73
Appendices	77
Appendix A.1	77
Appendix A.2	78
Appendix A.3	82
Appendix B.1	87
Appendix C.1.....	91
Appendix D.1	94
Appendix D.2	95

List of Acronyms

Acronyms

- HCI = Human-Computer Interaction
- VEs = Virtual Environments
- UMCG = University Medical Centre Groningen
- PACS = Picture Archiving and Communication System
- DICOM = Digital Imaging and Communications in Medicine
- CT = Computed Tomography
- MRI = Magnetic Resonance Imaging
- POI = Point Of Interest
- ROI = Region Of Interest
- CAS = Computer Assisted Surgery
- WIMP = Windows Icons Menus Pointers

Chapter 1.

Introduction

For many years medical two and three-dimensional images have been navigated and manipulated with mouse and keyboard to pan, zoom and change contrasts to get a clearer view of the patient's condition. However, there is a rising interest in other, more 'natural' interaction devices, which has mainly been triggered by the gaming industry. One well-known example is the Wii remote, which serves as a 'Motion controller' of interactive games and other applications. Another household gaming innovation was presented in 2010 by Microsoft when it revealed the 'Kinect'; a device that does not require extra peripherals (unlike the Wii) to determine parameters such as depth via its infrared sensor. The Kinect and other similar devices have caused a shift in the way people think about human-computer interaction from a traditional mouse and keyboard perspective to a more natural way of interacting by using gestures.

Consequently industries and researchers are increasingly interested in incorporating gesture-based interaction techniques in their products and services to create a more natural way of interaction and to enhance the user experience. One interesting example is the television industry, which is currently developing televisions with integrated cameras that aim to make the physical remote control obsolete. Another interesting area of innovation, which forms the basis of this research, is the medical imaging sector in which computed tomography (CT) and magnetic resonance imaging (MRI) studies are viewed in clinical image viewers and are traditionally controlled by mouse and keyboard.

1.1 Problem description

In recent years it has become common practice for surgeons to access and view patient's medical imaging studies before and during surgical procedures. Surgeons often visualize and manipulate these images on large monitors in the operating room in order to assist them during a surgical procedure, thus replacing the old-fashioned way of holding up analogue films against a light box. Currently, special software is used to access (a secure local server) to download a patient's image studies, and to display this information in a clinical image viewer. These viewers have certain functions such as scrolling through the images of a selected study, alter zoom-level and contrast values, but also more advanced functions such as measuring angles and line segments. Such a system is mostly used pre-operatively to refresh the surgeon's memory, and to display the most important information during the surgical procedure. Although this is a major improvement with respect to the analogue era, very few systems have been designed to allow for more practical and efficient exploration of these images during the actual surgical procedure, where time and sterile conditions are crucial for a successful surgical procedure.

Currently when surgeons want to get a better look at the images of the patient during surgery, he or she has two options. First the surgeon can ask an assistant to do this, which is time-consuming, distracting and may lead to errors due to the indirect form of communication. Secondly the surgeon can decide to interact with the computer him-or-herself, but this implies changing gloves each time the computer has to be operated. This again interrupts the workflow, costing precious time and may even endanger sterility (Schultz, Gill, Zubairi, Huber, & Gordin, 2003).

During an observational study by Grätzel et al. (2004) on the implementation of a non-contact mouse for surgeon computer interaction, a striking scene was observed in which a surgeon instructing an assistant took about 7 minutes to direct the assistant to click on the exact point of interest on the medical image. This example illustrates the problem at hand and calls for a solution that benefits the surgeon and indirectly the patient's situation.

Such a solution could possibly be found in gesture-based interaction techniques as mentioned earlier. The main benefit of a gesture-based interaction technique in the

operating room is that surgeons can directly interact with a clinical image viewer while operating on the patient. Such a system could potentially enhance the surgeon's level of control and thus might save precious time while maintaining a sterile environment.

1.3 This thesis

This thesis looks into the possibilities of gesture-based interaction to serve as a more natural and usable way of viewing and manipulating medical images as opposed to asking an assistant to control the traditional keyboard and mouse. In this case the aforementioned Kinect is used as input device due to its popularity, low-cost and many on-going developments in the medical imaging domain incorporating the Kinect in their products (see chapter 2.2).

Due to the lack of usability research conducted on gesture-based interaction techniques in the operating room (see chapter 2.3), we conducted a thorough usability study conducted to find out whether surgeons want this type of interaction, if it is suited for real operating room usage and how it can possibly be improved.

The obtained results give important insights in; the requirements of surgeons wanting to explore medical images during surgical procedures and the performance measures needed to evaluate such systems, if current state-of-the-art devices such as the Kinect meet these requirements and performance measures, and finally how the usability can be improved.

1.4 Research question and objectives

The main research question that is addressed in this thesis is:

How can a gesture-based interaction technique using the Kinect be implemented for operating room purposes?

This broad research question is broken down in the following sub questions:

- *What do surgeons expect from gesture-based interaction techniques?*
- *Can the Kinect serve as a better way of interacting with medical images in the operating room as opposed to asking an assistant?*
- *How can gesture-based interaction in the operating room be improved?*

In order to address these questions, the research is broken down into the following research objectives:

- **Explore the possibilities of gesture-based interaction techniques for operating room purposes.** Find out how surgeons regard gesture-based interaction techniques in the operating room, and identify possible requirements, restrictions and performance measures needed to evaluate such a system.
- **Evaluate a test-case system using the Kinect in an operating room environment.** Test a suitable gesture-based interaction system in a realistic setting with actual surgeons on prototypical tasks, and compare these results to a control condition in which the participant has to instruct an assistant on the same tasks.
- **Evaluate the performance of gesture-based selection techniques.** Find out which gesture-based selection technique is most appropriate for operating room purposes, due to the importance of accurate selection in medical images
- **Suggest usability improvements and guidelines for future systems.** Finally suggest improvements for the tested system and group all results in a set of guidelines for gesture-based interaction systems in the operating room.

1.5 Thesis organization

This thesis is organized into the following nine chapters:

- **Chapter 1, Introduction** - presents the motivation and research objectives of the thesis.
- **Chapter 2, Theoretical background** - reviews research literature relevant for surgeon-computer interaction and related work on gesture-based interaction in the operating room.
- **Chapter 3, Exploratory study** - describes interviews, operating room observations and questionnaire results concerning gesture-based interaction preferences.
- **Chapter 4, Practical background** - describes relevant practical information for the test case usability evaluation in Chapter 5.
- **Chapter 5, Test case usability evaluation** - describes a user-based usability evaluation of a test-case gesture-based interaction system using the Kinect.
- **Chapter 6, Performance evaluation of gesture-based selection techniques** - describes an experiment on accuracy measures of different gesture selection techniques and differing interaction zone areas.
- **Chapter 7, Usability recommendations** – elaborates on the results of the experiments in Chapter 4,5 and 6 in order to address usability improvements and guidelines for future gesture-based systems in the OR.
- **Chapter 8, Discussion** – discusses the results and limitations of this research and poses suggestions for future research.
- **Chapter 9, Conclusion** – answers the main research question.

Chapter 2.

Theoretical background

This chapter provides an overview of the literature relevant to understanding the domains of surgeon-computer interaction. The theoretical background presented is divided into three categories: a brief overview of the human-computer interaction domain and usability methods, gesture-based interaction and the Kinect, and finally related work on relevant gesture-based usability studies and gesture-based interaction in the operating room.

2.1 Human-computer interaction and usability

The Human-Computer Interaction (HCI) research area is often regarded as the intersection of social and engineering sciences in addition to design. The following definition by Bongers (2004) gives a good description of the field; “Human-Computer Interaction can be defined as the research field that studies, and develops solutions for, the relationship between humans and the technological environment”.

The main long-term goal of the HCI research area is to design systems that minimize the barrier between the user’s goals and the computer’s ‘understanding’ of these goals. This barrier is addressed by Norman’s Action Cycle (Norman, 1991), which describes how humans form goals and then develop a series of steps required to achieve these goals by using the relevant system. In this action cycle two types of mismatches might occur. The first is the ‘gulf of execution’, which describes the gap between the user’s perceived execution actions (or mental model) and the actual required actions of the system which is operated. Secondly the ‘gulf of evaluation’ describes the psychological gap that has to be crossed between the information representation of the system and the interpretation by the user.

In the operating room surgeons represent the users and the system is often a desktop computer. The interface that connects the system with the surgeon’s goals is in this case a clinical image viewer. In the current situation a surgeon has to control mouse and keyboard in order to view and manipulate medical images of the patient. The surgeon can also ask an assistant to do this for him or her during surgery, which turns the assistant into an indirect controller. This thesis explores the possibilities of gesture-based interaction to serve as a more natural, and direct form of interaction between user and system. This new form of interaction could potentially minimize the gulf of execution because of its designated naturalness by letting users communicate their intentions through gesturing to the system, which subsequently interprets these gestures as commands and actions.

Several usability methods exist to evaluate a product or technique on its efficiency and user satisfaction. The term usability is a broad concept, and generally refers to the ease of use and learnability of human-made objects. According to the ISO 9241-11 definition, usability is “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (ISO, 1998). More specifically in the HCI field usability refers to the attributes of the user interface that makes the product easy to use. According to usability pioneer Nielsen (1993) usability is not very well expressed in a definition but its concept is clearly reflected by learnability, efficiency, memorability, errors and user satisfaction. A system is said to be usable when it is easily learned by novice users, delivers high productivity, is easy to remember over time, has a low error rate and is considered pleasant to use.

The usability of a certain product can be evaluated by numerous methods; these methods can generally be divided into three separate categories (Dumas, 2003); *inspection-based*, *model-based* and *user-based evaluation*, of which the last category will be of main interest for this thesis.

Inspection-based usability evaluation is concerned with methods in which the evaluator inspects a system on its usability based on a series of heuristics or other predefined method. One great advantage of these methods is that it does not require any users, which often makes them very time and cost efficient. These methods have drawbacks for this study however. First of all they should be applied by multiple usability experts for it to be maximally effective (J. Nielsen & Landauer, 1993). Secondly specific domain knowledge is needed for usability experts to correctly evaluate the system at hand, in this case specific domain knowledge is needed of medical-image viewers and its use in the operating room, which is presumably not the average expertise of the average usability expert. And lastly inspection-based evaluations do not take performance measures of the users on the system into account, which in this case is very important considering the stress and time constraints in operating room environments.

Model-based usability evaluation methods are concerned with computational models of human behaviour and cognitive processes of how users would perform a certain task. These methods greatly differ with the other two methods in that they try to predict and describe a user's behaviour, whereas user-based methods can only retrospectively generate usability issues after a user has performed certain tasks and inspection-based methods can only guess on usability issues based on the heuristics used by and the knowledge and experience of the evaluator. For this study model-based evaluations also have certain drawbacks; first of all most models only model expert-user behaviour and thus cannot model novice-usage of the evaluated system. More importantly, model-based evaluations have mainly been applied to systems in which keyboard and mouse are used as controllers. Hardly any studies have been conducted on gesture-based interaction besides the study of Holeis et al. (2007) who have looked at modeling advanced mobile-phone interaction with the Keystroke-level model. Due to this major limitation for natural user interfaces and the explorative nature of this research in wanting to find out usability as well as functionality requirements, model-based usability evaluation will not be taken into account.

User-based usability evaluation is concerned with gathering input from relevant users interacting with the system or interface of interest. This type of evaluation is particularly relevant for user-centered design. The most widely used method in this category is the questionnaire, which measures the users' subjective preference after using the relevant system. These questions can elicit qualitative (open questions) as well as quantitative data (closed questions and scale responses). One important and widely used usability scale used in user-based evaluations is the Software Usability Scale (SUS) (Brooke, 1996).

Questionnaires are often preceded by other usability measures such as scenario-based testing. In these tests several (prototypical) tasks are presented in the form of scenarios, which explain what the participant needs to do on the system, but not how it should be performed (Dumas, 2003). In such an evaluation study, a participant is often asked to conduct several tasks in a controlled environment, while being monitored on several performance measures. If the evaluation focuses on comparison of multiple systems, then participants have to repeat the same tasks in different experimental conditions. In order to evaluate how well the tested system supports high-level real-life goals, the representative user task scenarios should include more than simple atomic tasks; these scenarios should always include high-level cognitive, problem-solving tasks that are specific to the application domain (Bowman & Hodges, 1999, chap. 11). Furthermore it is also important that these evaluations are performed in a representative natural working environment, so that obtained experimental results can be generalized (Dumas & Redish, 1999).

Letting participants perform a standardized set of tasks, gives the evaluator control over the experimental variables, over which aspects of the system are tested and over the performance measures obtained. In user-based usability studies objective as well as subjective measurements can be obtained, such as efficiency (e.g. time on task),

effectiveness (e.g. number of errors and actions) and satisfaction (e.g. user ratings on designated usability scales such as the SUS).

User-based usability evaluation is very suitable for the current study due to its explorative nature; finding out the feasibility and usability of a new interaction technique in a specific working environment, and the need for quantitative as well as qualitative data to see whether this system is preferred over the current situation.

2.2 Gesture-based interaction

Gestures

Gestures and gesture-based interaction are terms that are increasingly encountered in the HCI domain. Gesture-based interaction is a broad term and can refer to gesture recognition on touch-screen surfaces, device-based gestures (shaking a portable music player to skip to the next song) and freehand gestures (waving towards the television to switch to the next channel). In fact every physical action essentially involves some sort of gesture. To distinguish between these different types of gestures it is important to consider the nature of the gesture that is used to reach the interaction goal.

In this thesis we will only focus on interactions issued with the users hands to articulate certain gestures and which are recognized by the system. This stands in contrast to gestures that are issued by means of a device, handheld or mouse for example, or other form of transducer, such as a keyboard. When referring to gestures in this thesis the following description of a gesture is intended: “A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed” (Kurtenbach & Hulteen, 1990). This description clearly stipulates the importance of the intent of the user and the actual movement of the hands. Consequently this type of interaction is far richer and thus more complicated than any other type of interaction. This is due to the high number of degrees of freedom of gestures in comparison to two-dimensional input devices such as the mouse.

There are several different types of gestures, which can be categorized in several different ways. The taxonomy proposed by Karam and Schraefel (2005) fits the gesture interaction in this thesis best. It describes five gesture classes of which the following three focus on the tasks that people would like to accomplish with a gesture-based computer interface.

- *Deictic gestures*: these are manual pointing gestures, which are used to direct attention to specific events or objects in the environment. Bolt (1980) was early to note the intuitiveness and potential of deictic gesturing in gesture-based interfaces by letting users point at targets and select or manipulate objects by speech commands. Furthermore deictic gestures are often used in virtual reality applications for example (Zimmerman, Smith, Paradiso, Allport, & Gershenfeld, 1995).
- *Manipulative gestures*: these are gestures that are tightly mapped to the movement of a virtual object in the interface. These gestures can be performed on a surface or in mid-air and are sometimes accompanied by tangible objects such as a study in which a MRI scan is controlled by rotating a doll's head (Hinckley, Pausch, Proffitt, & Kassell, 1998).
- *Gesticulation*: concerns gestures that accompany everyday speech and are thus considered as one of the most natural forms of gesturing. Gesticulations are spontaneous and idiosyncratic movements of a user's hands during speech, which come naturally and thus do not require the user to learn these gestures (Karam & Schraefel, 2005).

Especially the first two gesture types are important for interacting with gesture-based systems. Deictic gestures are often used as a way to navigate the cursor, similar to the use

of a mouse but instead the users immediately points at the point of interest whereas the mouse often trails behind. A technique to implement deictic gestures in gesture-based interfaces is ray-casting for example (Vogel & Balakrishnan, 2005). In ray-casting the cursor is placed at the point where a virtual ray extends from the index finger intersects with the display.

Manipulative gestures are interesting for gesture-based interaction because they are often associated with movements made in real-life and thus often correspond to the mental model of the user. Rotation is such a gesture; people make rotating gestures when communicating the concept of rotation to each other, which is again tightly coupled to the real action of turning a physical object in several dimensions.

Gesture interaction challenges

When considering gesture-detection there are several major challenges. Namely when does a gesture start? Which gesture is being issued? And when does it end? These questions fundamentally distinguish this type of continuous interaction from device-based interaction, which is always a discrete form of interaction. Traditional input devices such as mouse and keyboard have one great advantage from a system's perspective: it is immediately clear when a user has issued a command. This is not the case with gesture tracking devices such as the Kinect. As long as the user is in range of the camera, a continuous input stream is generated of which it is hard to distinguish meaningful behaviour from irrelevant user behaviour.

From a user's perspective, this can sometimes lead to the so-called 'immersion syndrome', (Baudel & Beaudouin-Lafon, 1992) which refers to the unintended or inadvertent triggering of actions. This occurs when seemingly random movements are classified by the system as meaningful gestures, which subsequently trigger unwanted actions onscreen, frustrating the user. From a systems perspective, this phenomena is often referred to as 'gesture-spotting' in the literature (Lee & Kim, 1998). Gesture spotting has two major difficulties, namely segmentation and spatio-temporal variability. Segmentation refers to how the start and end of a gesture is determined by the algorithm, while spatio-temporal variability refers to the extent to which gestures vary dynamically in shape and duration. Several algorithms exist which have found ways of dealing with these problems, of which so-called Hidden Markov Models (HMM) are often used because this technique represents gesture-patterns as well as non-gesture-patterns but it also reflects spatio-temporal variability very well.

Gesture-based selection

Direct manipulation remains the current trend in user interface design, whether it is using a finger to touch and select an item on a tablet computer or pointing at a large screen with a laser pointer to select certain features. Selection plays an important role in the way users can achieve their goals with an interface. Selection tasks pose a real challenge for gesture-based interaction techniques, which also plays an important role in surgeon-computer interaction: surgeons not only select menu-items and certain patient studies but also conduct precise measurements on these images.

Although gesture-based interaction interfaces offer interesting possibilities, the simplicity and self-revealing nature of a WIMP- (Window, Icon, Menu, Pointer) style interface is a hard to ignore property, which should not be overlooked when designing a gesture-based interface. Furthermore gesture-based selection techniques have two disadvantages compared to a mouse or any other device-based interaction technique, namely the absence of a clear affordance for selection and the absence of physical feedback both due to lack of physical buttons.

Bowman and Hodges (1999) describe that most of the interactions in a virtual environment fall into one of three categories: selection, manipulation and navigation. To a certain extent this is also the case for natural user interfaces, with an emphasis on selection and manipulation. One popular solution for selection in gesture-based

interfaces is to use dwell time thresholds, also referred to as 'Hover', which activates a select command when a user points at a particular target area for a predefined amount of time. This can be achieved with an extended finger (Vogel & Balakrishnan, 2004) but also with an eye-tracker for gaze estimation for example (Zhang & Mackenzie, 2007). Another popular selection method is 'Push' which requires the user to stretch his dominant or non-dominant hand towards the screen until a predefined threshold on the z-axis is reached, after which the target is selected. One limitation of these techniques however is that there is no kinaesthetic feedback confirming the click action to the user.

Vogel and Balakrishnan (2005) argue that the hand can serve as its own source of physical feedback, also referred to as kinaesthetic feedback. This finding was implemented by Grossman et al. (2004) in a technique called 'Thumbtrigger', in which the hand is shaped like a pistol. Clicking is done by pressing the thumb on the bent middle finger, pretending to press a button. This technique is still in an experimental phase in which a special glove with sensors is required to register this clicking gesture, but in the future this should also be able to be detected with high-resolution depth-cameras.

The Kinect

Currently one of the most popular gesture-based interaction devices is the aforementioned Kinect. The Kinect (see Figure 1.) is a camera peripheral by Microsoft initially developed for the Xbox 360 video game console¹ in 2009, while later in 2012 a Kinect for Windows was released².

The Kinect captures the user's movements and translates these into commands for the console. The Kinect 'understands' gestures and spoken commands. It is the world's first device to combine full-body three-dimensional motion capture, facial and voice-recognition and dedicated software for consumers as well as a SDK for developers.

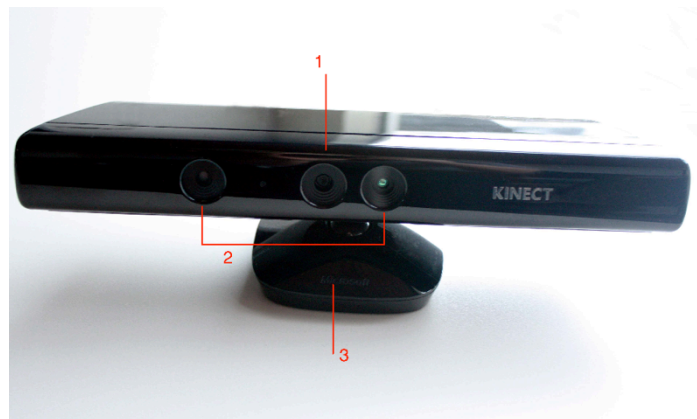


Figure 1. Xbox 360 Kinect sensor. 1. Indicates the built-in RGB camera. 2. 3D depth sensor consisting of an infrared laser projector left and monochrome CMOS sensor right. 3. Indicates the motorized tilt stand.

The Kinect device consists of a RGB camera, 3D depth sensors, a multi-array microphone and a motorized unit used to alter tilt level of the Kinect. The RGB camera used for creating a video stream of the user and for enabling facial recognition can process three basic colours and generates a video output at a frame rate of 30 Hz and a maximum resolution of 640x480 pixels in 32-bit colour, similar to many commercially available webcams. The 3D depth sensor consists of an infrared laser projector, which emits infrared light that passes through a diffraction grating, and a monochrome CMOS sensor,

¹ <http://www.xbox.com/en-US/kinect>

² <http://www/microsoft.com/en-us/kinectforwindows>

which can detect the reflected infrared lighting. The relative geometry between the IR projector and the camera as well as the projected IR pattern are known, after which a three-dimensional reconstruction can be calculated by using triangulation (Khoshelham & Elberink, 2012). The multi-array microphone consists of four separate microphones, processing 16-bit audio at a sampling rate of 16 kHz, enabling voice recognition. Furthermore it is capable of recognizing different users and distinguishing between background noise and meaningful voice-commands. The Kinect features a horizontal field of view of 56 degrees and a vertical field of view of 43 degrees and the operating distance ranges between 0.8 m and 3.5 m. Within 2 m usage the spatial resolution in the X/Y plane is 3 mm and 10 mm for the Z plane. Finally this produces a data stream with a resolution of 640x480 pixels at 30 fps. For a detailed description of the algorithms used for gesture-recognition, the reader is referred to an excellent article by Shotton et al. (2011).

2.3 Related work

Usability research on gesture-based interaction

Although there is a relatively large amount of usability research data on classical user interfaces and on suitable usability methods (Hornbæk, 2006), there is hardly any usability research conducted on gesture-based interfaces. Usability experts Donald Norman and Jakob Nielsen (2010) point out that in the “rush to develop gestural interfaces – ‘natural’ they are sometimes called - well-tested and understood standards of interaction design were being overthrown, ignored and violated”. They go on to acknowledge that gestural systems require novel interaction techniques, “but these interaction styles are still in their infancy, so it is only natural to expect that a great deal of exploration and study still needs to be done”. They conclude by stating “we urgently need to return to our basics, developing usability guidelines for these systems that are based upon solid principles of interaction design, not on the whims of the company human interface guidelines and arbitrary ideas of developers”.

Despite the lack of general usability guidelines for natural user interaction, there are certain relevant areas that are looking into the usability of certain aspects of natural user interfaces; such as the virtual-environments research area, as well as ‘intuitive’ gesture studies to determine the most optimal gestures for certain functionality, and more elementary research on gesture-based selection techniques. These areas are shortly discussed below.

Virtual-environments (VEs) are computer-simulated environments in which the user is immersed and can often interact with, often known as ‘virtual reality’. These simulations are often displayed on stereoscopic displays in cave-like environments or special head-mounted glasses. Bowman and Hodges (1999) observe rapid advances in display technologies, graphics processors and tracking systems but a lack of knowledge on complex interaction in such environments: “there seems to be, in general, little understanding of human-computer interaction ... in three dimensions, and a lack of knowledge regarding the effectiveness of interaction in VEs”. To this extent their paper describes a methodology with which usability of interactively complex VE applications can be improved. This framework stipulates the importance of formalizing interaction characterizations into taxonomies for an overview of essential functionality, listing ‘outside’ factors that might influence task performance, listing multiple performance metrics for VE interaction tasks, and methods for measuring performance. Finally quantitative and general experimental analyses should be developed in order to compare the performance of different interaction techniques on the universal tasks described in the aforementioned taxonomy.

Several studies have been conducted to find out which gestures best suit the application from a user’s perspective as opposed to a system’s perspective. Nielsen has proposed a procedure for finding the most intuitive and ergonomic gestures for a certain interface (Nielsen, Störning, Moeslund, & Granum, 2004). One example of intuitive gesture

research for HCI has been conducted by Radu-Daniel Vatavu (2011) who has looked at user-defined gestures for free-hand television control by conducting an agreement analysis on user-elicited gestures and found an average of 41.5% agreement on functionality. Another example is a study by Wim Fikkert in his PhD thesis on “gesture interaction at a distance” (Fikkert, 2010) who looked into which gestures are most suited for large display control from a short distance. In this study he uses a Wizard-of-Oz technique to elicit gestures from uninstructed users asked to issue certain commands through gesturing alone. Results indicated that these gestures were “influenced profoundly by WIMP-style interfaces and recent mainstream multi-touch interfaces”. These gestures were later tested and evaluated in a 3D and 2D prototype application and found results that comply with the abovementioned results about the intuitiveness of the chosen gestures and physical comfort levels. One interesting concluding remark by Fikkert is the observation that technological developments that reach the general public influences their perception of intuitiveness and naturalness, a notion that should be taken into account when developing natural user interfaces. More recently Nancel et al. (2011) looked into three key factors in the design of mid-air pan-and-zoom techniques: uni-manual vs. bimanual interaction, linear vs. circular gesture movements, and level of guidance (feedback). They found that bimanual interaction, linear gestures and a high-level of guidance produce the most optimal performance.

Research on gesture-based selection techniques

In the operating room it is important that surgeons are able to select targets accurately and easily when conducting measurements on a tumor for example. Most studies on the usability of gesture-based selection techniques involve high cost motion detection setups, often accompanied by special tracking markers. To my knowledge, no gesture-selection research has been conducted on low-cost, popular motion detection devices such as the Kinect (Chapter 6 is concerned with an evaluation study of gesture-selection techniques using the Kinect as input device).

In order to measure the effectiveness of the Kinect in its ability to select targets as fast, accurate and enjoyable as possible it is important to quantify the performance of several selection techniques. Two popular Kinect selection techniques include the aforementioned ‘Hover’ and ‘Push’ gesture. These gestures are easily detected by the Kinect and are often applied in interactive games, for which the Kinect was originally developed. One limitation of the Kinect is its resolution, which makes it very hard to detect separate fingers, which are required for precise selection techniques such as ‘AirTap’ and ‘ThumbTrigger’.

One often used standard for measuring performance is the ISO 9241, part 9 standard “Ergonomic requirements for office work with visual display terminals” (Iso, 1998). This standard establishes uniform guidelines and testing procedures for evaluating computer pointing devices. The performance measurement proposed in this standard is ‘throughput’. Throughput, in bits per second, is a combined measurement derived from speed and accuracy responses in a multi-directional point-and-select task (for a detailed account of throughput see section 6.2).

This testing standard has been applied by several studies. One study by Douglas, Kirkpatrick and Mackenzie (1999) evaluated the scientific validity and practicality of this standard by comparing two different pointing devices for laptop computers, a finger-controlled joystick and touchpad. Results indicate that a significant effect was found for throughput in the multi-directional task in favour of the joystick. A later study by Mackenzie, Kauppinen and Silfverberg (2001) conducted a similar study, but in this study several other properties besides throughput were also taken into account, such as ‘target re-entry’, ‘movement direction change’ and ‘movement offset’. These measures capture aspects of movement behaviour *during* a trial as opposed to the original throughput measurement, which only looks at performance *after* each trial. A more recent study by Zhang and Mackenzie (2007) was conducted in which the same ISO testing standard was applied to eye tracking selection techniques. In this study they looked at three different

selection techniques; the 'Eye-tracking short' technique in which the eye controlled cursor has to dwell on the target for 500 ms, the 'Eye-tracking long technique' with a dwell time of 750 ms and the 'Eye+Spacebar' technique in which participants have to point with eye movements and select by pressing the spacebar. The 'Eye+Spacebar' technique turned out to be the best of the three techniques with a throughput of 3.78 bits/s, which was closest to the mouse condition of 4.68 bits/s.

Research on gesture-based interaction in the operating room

The amount of (usability) research on gesture-based interaction is generally slim and only a few studies have been conducted that explore the possibilities of gesture-based interaction of medical images and patient information for operating room purposes. One of the first studies is the earlier mentioned study conducted by Graetzel et al. (2004) who explored the possibilities for a 'non-contact mouse for surgeon-computer interaction' by developing and testing a system that uses a colour stereo camera as input device and advanced image-processing software to detect movements and gestures. This system was tested on (limited) usability by letting 16 subjects test the interface on certain predefined tasks (of which some were timed), after which a questionnaire was filled out. Overall positive results and insights were obtained: all subjects were able to learn how to use the system fast; a majority of the subjects preferred the "push to click" selection mechanism; and subjects had difficulty working inside the three-dimensional workspace, but with experience they were able to gain access to all points on the display.

Limitations of the tested system are: it relied on good-lighting conditions; surgeons had to interact in a predefined interaction area and could not be tracked; it only focused on hands without tracking a surgeon as a whole and thus could not distinguish between several different hands, which is a desired feature in a crowded surgical environment.

Limitations of the used method are: only time on task and a single questionnaire were used as usability measurements, which is rather little; and of the 16 participants only two subjects were medically educated (medicine students), which hardly represents the actual user domain.

Wachs et al. (2008) looked at a two-dimensional camera for browsing medical images in a sterile medical environment in which they used a system that was user independent and could recognize gestures as well as postures. Whereas gestures imply a specific movement over time, postures imply a certain predefined static pose over time. Their system was tested by surgeons during a neurological, and the performance of the system was evaluated on gesture recognition accuracy, task completion time and number of excess gestures used for ten non experienced users and were afterwards queried on the ergonomic aspects such as comfort and intuitiveness. Results showed that the overall recognition accuracy of the system was 96 percent. Looking at the accuracy of the rotation gesture a mean absolute error of 3 degrees was measured. Furthermore when looking at the task completion time, a learning effect was observed which levelled off at the 10th trial. Finally the ergonomic aspects of the system revealed that subjects were moderately positive (5.8 on average out of 10) about the strength of the comfort level of the used gestures, and strongly positive (7.9 on average out of 10) about the intuitiveness of the used gestures.

Limitations of the tested system are: functions are activated with awkward gestures (for example: zoom mode is activated with a counter clockwise rotation of the wrist); some functions can only be activated while holding an instrument (such as rotation);

Limitations of the used method are: only 10 non-experienced (and not medically educated) users were used to test the system; the rotation accuracy was determined with 1 experienced user; and limited qualitative usability measures were obtained (only 2 questions about the interaction in general).

A slightly different technique was applied by Kipshagen et al. (2009), in which they made use of stereo-cameras to triangulate hand positions in 3D and map these to the 2D environment of the medical image viewing software. Another distinguishing feature of their system is that the cameras were mounted near the floor of the operating room and thus look upwards toward the ceiling for postures issued by the surgeon. This special

setup provided good results with and without artificial illumination. Only hand recognition results of the used algorithms are presented, of which a correlation of 0.94 was found between a manually segmented hand and a test set of 312 images of hands.

Limitations of the tested system are: the system essentially recognizes arbitrary postures which activates certain functionality (for example zooming is activated by pointing with one finger) after which the parameter of interest is increased by moving the hand forward from the baseline; furthermore because the system looks upward, the surgeon has to be in this designated field of view, and cannot move around (to the other side of the table for example).

Limitations of the used method are: hardly any usability evaluation was conducted, only the systems' pointing precision was validated by having volunteers point their index finger at a computer generated cross hair in a web-based commercial flash game.

One of the first studies using a commercially available depth-camera, in this case the Kinect (see 3.2.3 For a detailed description), is a study by Ebert et al. (2011). Similar to the aim of this thesis, they conducted a feasibility study in which 10 medical professionals were asked to re-create 12 images from a CT data set. Response times and usability of the system were compared to standard mouse/keyboard control. Participants required 1.4 times more time to recreate images in the gesture condition as opposed to the mouse/keyboard condition (75.1 seconds on average versus 52.1 seconds). Furthermore the system was rated 3.4 out of 5 for ease of use compared to mouse and keyboard.

Limitations of the tested system are: that it relied heavily on voice-commands for selecting different 'control modes' which can be problematic in the noisy operating theatre and as noted by the authors it can be difficult to recognize users with certain accents, in this case German accents. Additionally mainly constrained gestures are used, for example users can browse through the current patient study (only if the stack navigation mode is selected through a voice-command) by moving one hand up or down. So functionality must be explicitly selected after which moving the hands upward is similar to pressing the up key on a keyboard for example, this is different from 'natural user interfaces' in which each function is addressed by a specific (intuitive) gesture.

Limitations of the used method are: that the gesture-based system was compared to a control condition in which the surgeon directly interacts with mouse and keyboard, which is not a very realistic control condition. This is not a realistic control condition because it is to be expected that the surgeon is a lot faster with mouse and keyboard due to years of experience. A more realistic control condition would be a surgeon asking an assistant to interact with keyboard and mouse. Furthermore the used usability questionnaire was restricted to three questions concerning the general system (intuitiveness of use overall, accuracy of gesture control, and accuracy of voice control), which is fairly limited.

Very recently a new innovative concept was added to gesture-interaction in the OR by Bigdelou et al. (2012). They presented a touch-less, gesture-based interaction framework that lets surgeons define a personalized set of gestures instead of the predefined gestures in the system discussed above. This system does not rely on any cameras but uses several wireless and inertial sensors, placed on the arms of surgeons and thus eliminating the dependence of good lighting and the surgeon having to be in the line-of-sight. One challenge however is the importance of distinguishing between intended gesture commands and other movements of the arm. The authors therefore introduced voice-activation and a wireless handheld switch. A user study was conducted in which participants had to complete a training phase in which the gestures were personalised after which they were tested on a CT dataset on a well-defined test task. Time-on-task and accuracy were used as usability measures as well as a usability survey. This survey showed that participants were generally positive about the system, except for its speed when compared to mouse and keyboard control. Besides the survey however no detailed results of time-on-task were discussed besides the accuracy rates of 2 to 4% deviation of the parameter range, respectively depending on the voice-activation or the hand-held switch activation condition.

Limitations of the tested system are: that accidental interaction with the system forms a problem that was solved with either voice-activation or a handheld switch, which are both suboptimal solutions.

Limitations of the used method are: Only time on task was measured as performance measure of the system, but not the accuracy of the system for example.

Overall, the various studies on gesture-based interaction show that most participants were positive about its usefulness and practicality in the operating room. Most systems were able to capture the intended action, but all the methods used in the studies described above either did not incorporate usability measures, or only used limited usability measures. The different usability methods described above will be taken into account when thoroughly evaluating a possible test case system on its usability in Chapter 5. Important usability and performance measures include: questionnaires concerning the usability of the system as a whole, gesture specific questions, time on task, accuracy of recognition, and accuracy of selection.

Chapter 3.

Exploratory Study

In order to get a better idea of the surgeon's preferences concerning medical image interaction before and during a surgical procedure, this chapter describes interviews, operating room observations and questionnaires with medical professionals. This will give important insights into the requirements of a possible gesture-based system for operating room purposes and will provide guidelines for implementation and evaluation in Chapter 5.

3.1 Interviews and operating room observations

Before an effort is made to test a gesture-based interaction system in the operating room, it is interesting to find out whether surgeons are familiar with such systems and whether they need or in would welcome a gesture-controlled medical image viewer. To this end a semi-structured interview was conducted with two orthopaedic surgeons. Questions for example included asking them to describe the current operating room situation regarding patient image viewing, how they regard gesture-based interaction techniques for operating purposes and which functionality is most important and thus should be present in a novel system (see Appendix A.1 for the full interview).

The surgeons indicated that they always consult medical images before but also regularly during a surgical procedure. During surgery they mostly ask an assistant to do this for them, but complain that the result is never exactly how they envision it in their minds.

They also stated that they would like to access the images more often than at present due to the discussed limitations, and they thought a gesture-based system could probably solve these limitations. Most used functions of the current image viewer include basic functionality such as zooming, changing contrast, scrolling through an image series and conducting measurements. Both surgeons indicated that they would welcome a gesture-based alternative and clearly saw the benefits of such a system and thought it would be interesting to test a potential system in a more realistic scenario.

Both surgeons also raised the possibility to attend specific image-guided surgeries. Three orthopaedic surgical procedures were observed in total: a removal of an osteochondroma in the distal femur (18 year old male); an excochleation of a tumor in the proximal humerus (65 year old female); and a removal of a tumor in the femur using computer-assisted surgery (43 year old female). Especially the last procedure was image intensive and was performed using so-called "Computer Assisted Surgery" (CAS). This technique is used to assist the surgeon in presurgical planning and for guiding or performing surgical interventions. The main objective of this technique is to reconstruct a three-dimensional image of the patient's affected area, so that no extra intra-operative images have to be made, reducing the amount of radiation in the operating room. Another important feature of CAS is that it can be connected to medical instruments with wireless sensors, so that these can be tracked during the image-guided navigation process (the surgeon can see how far away he or she is from the tumor for example). The position of the instruments and other marker points are simultaneously shown in the three-dimensional reconstruction on the screen (see Figure 2. label 1).

During these surgical procedures on average eight surgical staff members are present, including surgeon, surgical assistant, two assistants (one reaching out instruments and one gathering supplies from outside the clean air flow), one resident, one to two anaesthetists and one intra-operative medical image specialist. Sterility is very important during surgery, especially in the area surrounding the patient, referred to as the 'clean-air flow zone' (see Figure 2. label 4 and 8). This area designates the most sterile area of the operating room and is projected by a laminar airflow system, which filters out small bacteria-laden particles from the air in the OR. This is the main area where a gesture-based interaction system would be very useful and allow for the most sterile form of interaction.



Figure 2. Several images of operating room observations. 1. System used for “computer assisted surgery”. 2. 40-inch wall-mounted monitor. 3. Computer used to edit and view patient information, which is also connected to the large monitor on the wall. 4. Vent, which is an important component of the clean airflow system. 5. Several large OR lamps. 6. Surgical assistant. 7. Leading surgeon. 8. Marker indicating clean airflow zone. 9. Field of view between surgeon and monitor with medical images of the patient.

The following information regards the use of patient images during surgery. Surgeons often select the patient images before surgery starts. After some basic editing functions such as altering contrast and zoom-level it is projected on the large 40-inch monitor mounted on the wall (see Figure 2. label 9). This is often a single study with one or two different viewports, which can be images viewed from different longitudinal directions for example. These studies are often left unchanged during the observed surgical procedures, except for the abovementioned computer assisted surgery procedure. Due to the large size of the monitor it is most often possible to see the images during surgery, but it does not always mean it is in the line of sight of a surgeon.

Other interesting observations include the presence of two-to-three pedals in the vicinity of the patient. Also the surgeon does not always stand while operating on the patient, sometimes he or she sits. Furthermore there is a large amount of medical equipment present in and around the clean-air flow zone, making it very difficult for a surgeon to move around. For a schematic overview of the operating environment see Figure 3.

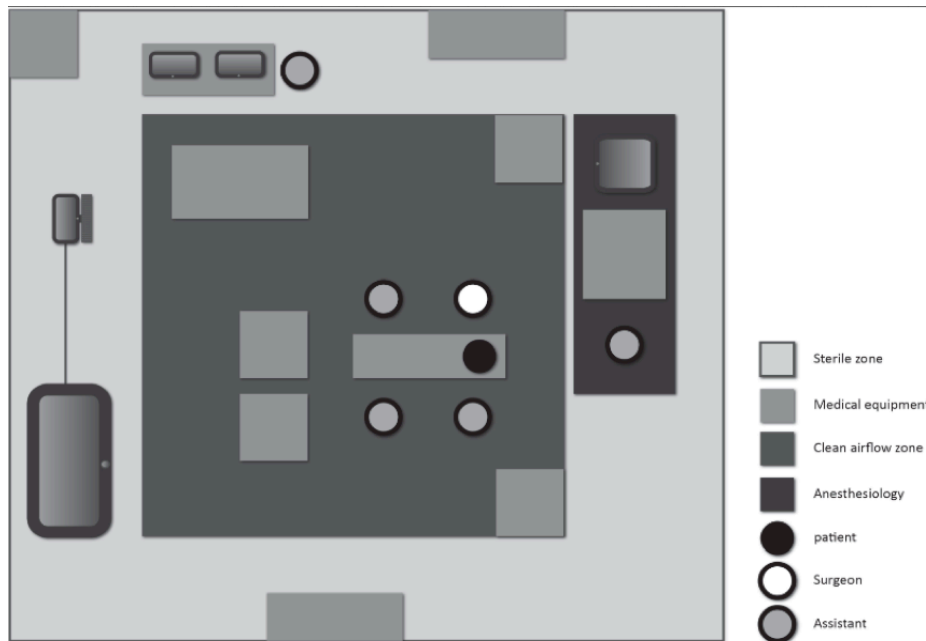


Figure 3. A schematic overview of the operating room environment, depicting the relevant sterile areas, equipment and personnel.

3.2 Online questionnaire

The former interview with orthopaedic surgeons and operating room observations were both indicative of the need for gesture-based interaction in the operating environment. To find out whether the chosen test case system meets the requirements of surgeons in general and how such a system should be implemented and evaluated, an online questionnaire was created and distributed under as much surgical staff in the hospital as possible.

Thesistools¹ was used to create the online questionnaire, which can be found in Appendix A.2. The questionnaire consisted of nine multiple-choice and four open questions and was accompanied by an introduction and a video² displaying the possibilities of gesture-based interaction for operating room purposes so that every respondent was informed about this new field of research. A web link to this questionnaire, accompanied by an email shortly describing the purpose and importance of this questionnaire, was emailed to five heads of departments and distributed among their surgical staff members.

Twenty-four respondents filled out the online questionnaire. The majority of respondents were specialized in general surgery (N=9). Other respondents were specialised in: orthopaedic surgery (N=5), vascular surgery (N=3), trauma surgery (N=3), oncology (N=2), urology (N=1) and abdominal surgery (N=1).

The most interesting results are summarized and discussed below; (a complete overview of the results can be found in Appendix A.3.)

- Generally respondents were already familiar with gesture-based interaction: 10 had already experienced gesture-interaction; 8 had heard of it; and 7 were unfamiliar with the concept.
- Hardly anyone was familiar with the possibilities of gesture-based interaction in the operating room (1 had already tested it once, 6 had heard of it while 17 were unfamiliar with the concept).
- The most frequently mentioned functionality of the medical image viewer in order of importance: animating through a patient study (N=24), zooming in and out (N=21), conducting measurements (N=18), adjusting window-level (N=16),

¹ <http://www.thesistools.com/>

² <http://www.youtube.com/watch?v=CsIK8D4RLtY>

translation (N=16), pointing (N=15), selecting different patient study (N=12) and finally rotation (N=10). The following functionality was considered to be used less frequently however: cutting out region-of-interest (N=3) and marking region-of-interest (N=3)

- When asked how often the respondent has to disrupt his or her surgical procedures, this appears to be about 0.78 times per procedure on average (N=20). Whereas he or she has to indirectly ask an assistant to do this 1.51 times on average (N=18).
- The majority (N=20) of respondents would want to interact with patient images more often than currently, in an ideal situation in which the workflow is not disturbed by either sterilization procedures or time loss due to indirect communication with an assistant.
- For activation and deactivation purposes of the system, 'voice-activation' was the most popular method (N=8), while others (N=6) preferred pointing at a special activation button on screen or (N=6) assuming a distinguishing posture. Footswitch activation (N=2) or asking an assistant (N=1) were less preferred options.
- Most respondents indicated that they would prefer to use one-handed gestures (N=15) if this were possible.
- For the implementation of a gesture-based system, respondents would like to use a moveable arm (similar to the OR lamps) with attached monitor (N=11) or a combination of the current large wall-mounted monitor and separate monitor to interact with (N=9). Only a few respondents (N=4) would want to use the current setup, with only the large wall mounted monitor.
- When asked about who should be able to control the system, (N=13) respondents think that the surgeon and assistants should only be able to control a gesture-based system, whereas (N=9) respondents think that everyone in the OR should be able to control it.
- When asked if the respondent regards a gesture-based interaction technique as a promising alternative for the current OR situation, the majority (N=17) responded positively, the remaining respondents (N=7) responded 'maybe'.

At the end of the questionnaire, respondents were free to leave any remarks or thoughts on gesture-based interaction. Remarks included "I don't know if large gestures are convenient in the OR", "Gesture controlled interactions can interrupt the laminar airflow, which is not always desirable", "Hand and arm movements should be kept small to ensure asepsis, nice idea", "Nice idea! This should be pursued!".

Both interviews as well as operating room observations indicate the need of surgeons for a more direct, sterile and easy to use technique for interacting with a patient's medical images. The interviewed and observed surgeons were very positive about a possible gesture-based system, and offered constructive remarks as well as the willingness to participate in a possible usability evaluation. Furthermore the hospital-wide distributed questionnaire indicated similar results and useful user preferences and insights for a gesture-based system.

Above results indicate that gesture-based interaction in the operating room is considered most welcome and worthwhile testing more thoroughly in a controlled operating room environment, which will be pursued in chapter 5.

Chapter 4.

Practical background

Due to the insights and preferences obtained in the previous chapter, a suitable gesture-based system was found that met these requirements. This chapter provides a description of the relevant software used in this thesis. First the current clinical image viewer is described after which the test case gesture-based system is described.

4.1 Current clinical medical image viewer

The medical image viewer currently used in the operating theatres in the UMCG is the PoliPlus Web1000 viewer (see Figure 4.); a web-based patient image viewer connected to the hospital-wide PACS (Van Ooijen, 2005), which stands for Picture Archiving and Communication System. DICOM (Digital Imaging and Communications in Medicine) formatted patient images are retrieved from the PACS database and loaded in the clinical web viewer, in which the surgeon can view and edit these images.

Common imaging modalities include CT (Computed Tomography) and MR (Magnetic Resonance) images. A set of images acquired from a particular patient during one scanning procedure is referred to as a study, whereas a single image or a sequence of 'slices' is called a series.

Very often surgeons and other healthcare professionals compare a new study with an old study of the same patient. The most common actions on studies include scrolling through the images of a series, manipulating images (e.g. adjusting zoom-level and contrast) and performing measurements (e.g. length measurements of a tumor).

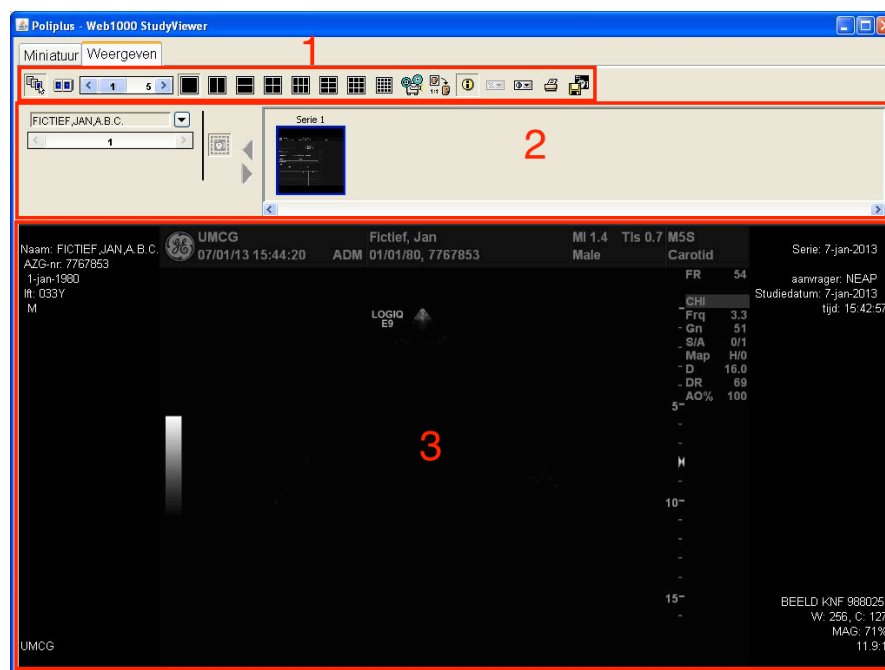


Figure 4. User interface of the PoliPlus Web1000 viewer. Frame 1 contains a toolbar with important functions such as selecting the amount of viewports, changing contrast and toggling patient info on and off. Frame 2 contains worklist filters in which the surgeon can select the relevant patient, study and from a thumbnail preview of contained image series. Frame 3 displays the selected series and important patient information.

4.2 Gesture-based medical image viewer

The clinical viewer chosen for the usability evaluation in chapter 5 is the open-source MITO (Medical Imaging Toolkit) PACS-integrated medical image viewer developed by the iHealthlab at the Institute for High Performance Computing and Networking in Italy¹. The developers describe MITO on their website as:

“The "MITO - Medical Imaging TOolkit" project coagulates a number of activities aimed at defining and implementing an open-source, cross-platform software architecture for advanced Medical Imaging. MITO toolkit makes it possible to fetch radiological information and images stored in a PACS according to the standard format DICOM, then provides the final user with basic functionalities such as 2D-3D visualization (VR, SR, MIP), image segmentation and fusion, ROI. Moreover, MITO provides interaction techniques for manipulating 3D medical data in a virtual environment by 2 DOF input devices.”

The MITO interface resembles that of the PoliPlus Web1000 viewer and DICOM viewers in general (see Figure 4. and Figure 5.). We used this system because the open-source software is free to use and its GNU General Public License makes it possible to alter the source code and freely distribute versions while only having to credit the original authors. But the most important reason is that it uses the Kinect as input device, which will be discussed below. To find out whether healthcare professionals regard MITO with the Kinect suitable as input device, a demo was given and an interview was conducted with two surgeons (see section 3.1). Also an online questionnaire, considering the preferences of medical image interaction during surgery, was distributed hospital-wide to see whether the preferences of healthcare professionals matched with the capabilities of MITO (see section 3.2). Finally the team of researchers and developers behind MITO and the Kinect lead by Luigi Gallo are very interested in gesture-based interaction and medical visualization². Gallo and his research group were informed about this research project and were very interested and willing to provide (technical) support throughout the project.

¹ <http://ihealthlab.icar.cnr.it/index.php/projects/9-mito.html>

² <http://www.luigigallo.net/publications>

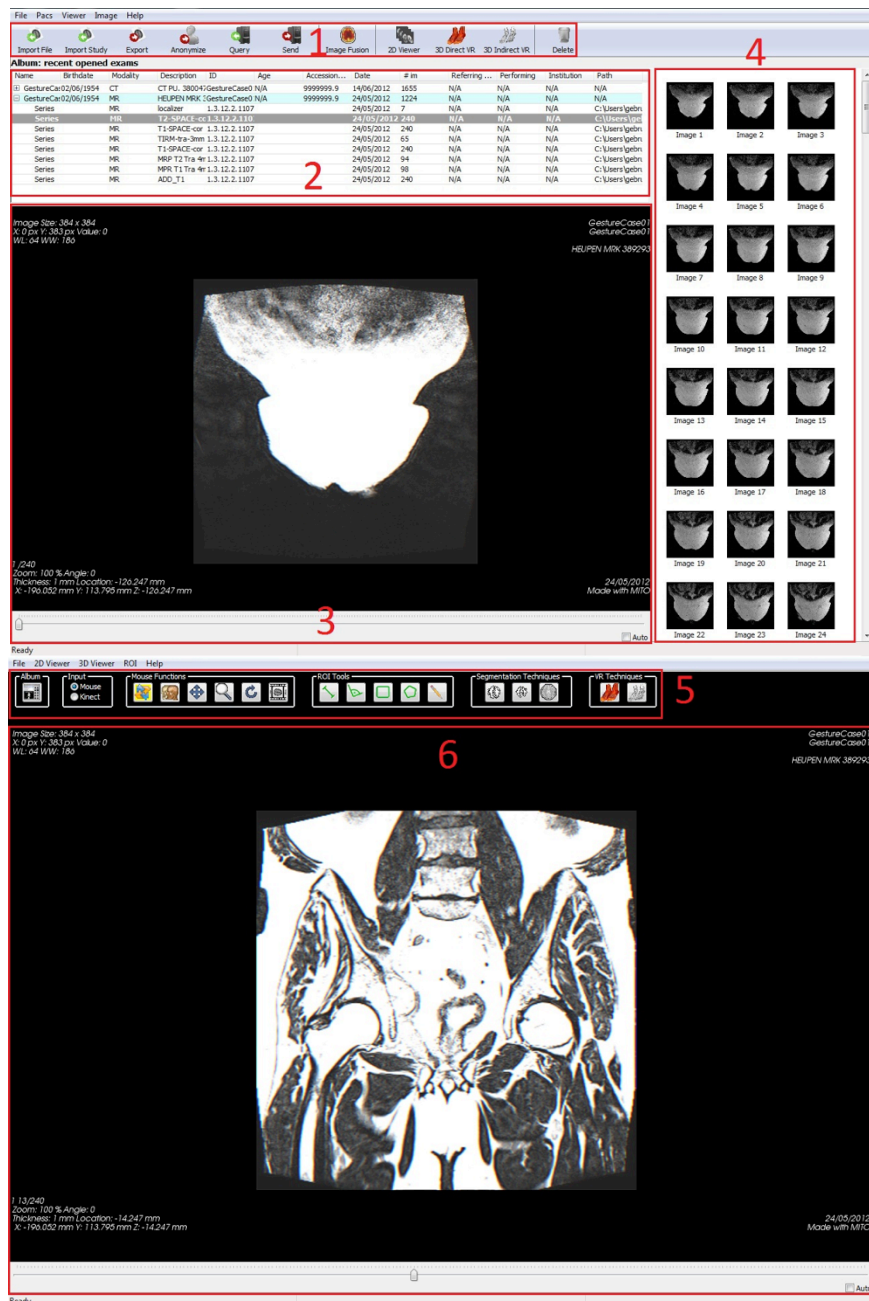
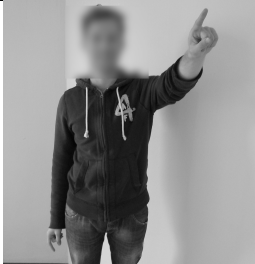
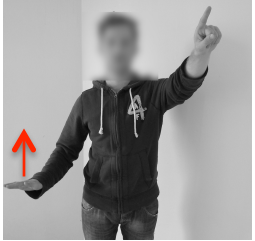





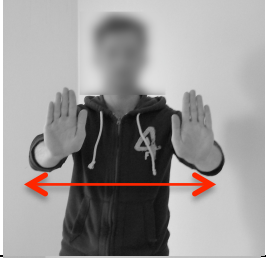


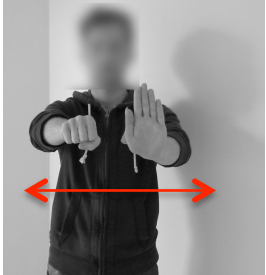

Figure 5. User interface of the tested MITO DICOM viewer, which consists of two different screens. The first screen is used for selecting the right patient studies whereas the second screen is used to explore and manipulate a selected image series. Frame 1 contains the toolbar for the selection screen. Frame 2 displays all studies and series, while frame 3 and 4 both display previews of a selection. Frame 5 in the second screen contains a toolbar for editing the selected study and frame 6 displays the actual series. Notably only frame 6 is displayed when the Kinect is used as input device.

As mentioned earlier, the same researchers at the iHealthLab in Italy have developed a plugin enabling the viewer to be controlled by Microsoft's Kinect camera (Gallo, Placitelli, & Ciampi, 2011). This plugin allows users to execute basic tasks such as animating (scrolling) through images of a patient study, zooming, translating, rotating and pointing, and more complex tasks such as changing window-level (contrast), conducting measurements and selection a region-of-interest (ROI). Recognized gestures contain both static elements, i.e. postures, and dynamic elements, i.e. gestures. The postures are used to discriminate between possible functions, while gestures define the parameters of the

‘selected’ functions. For example, when a user wants to zoom in on an image, he or she has to assume the zoom posture (keeping two hands opened together), while altering the distance between both hands determines the zoom level, similar to the popular ‘pinch-and-zoom’ gesture on touchscreen devices. One important aspect of this interface is that it relies on an ‘activation area’. This is a predefined area in front of the person being tracked, and checks whether at least one hand lies within it. More specifically this is more than 55% of an arm’s length, only then will the system try to recognize any meaningful gestures. This is implemented to ensure that no accidental movements are detected as gestures, and it also allows the system to minimize unwanted state transitions between gestures. Finally the system depends on a calibration procedure, which takes the user less than 30 seconds according to Gallo et al. This procedure is implemented in order to determine the user’s dominant hand, to map the interaction space of the user, to tune the parameters of the filters used in the underlying algorithms and finally to compute the surface of the palm. This last point is important because the system can distinguish an opened or closed hand, and this property is used to discriminate between several different functions (For an overview of MITO functionality see Table 1).

Table 1. Overview of MITO’s gesture functionality. The system can distinguish gestures based on dominant/non-dominant hand usage in the ‘activation area’ and whether these hands are open or closed.

Function/state	Static posture recognition	Dynamic gesture recognition	Illustration
Point	Only the dominant hand is active	Movements in the XY plane	
Select	Both hands are active, while already in the pointing state	Wave the palm of the non-dominant hand up and down in the activation area	
Clear selection	Only the dominant hand is active	Waving/Cleaning movement for 500ms	
Extract ROI	Folded/crossed arms	Movements in the Z plane	

Animate	Only the non-dominant hand is active	Movements in the X plane		
Zoom	Both hands are active and both palms are faced forward	Discordant movements in the XY plane		
Translate	Both hands are active and both palms are faced forward	Concordant movements in the XY plane		
Change window-level	Both hands are active with one opened palm faced forward and one clenched fist	Concordant movements in the X plane		
Change window-width	Both hands are active with one opened palm faced forward and one clenched fist	Concordant movements in the Y plane		
Rotate	Both hands are active with clenched fists	Movements in the XY plane		

Chapter 5.

Test case usability evaluation

Due to the positive results of the exploratory study in Chapter 3 concerning the preferences of surgeons of medical image interaction in the operating room and the complying and promising properties of the described gesture-based interaction system in Chapter 4, it was time to subject this system to an extensive usability evaluation to test whether gesture-based interaction could really be implemented in the operating room.

This chapter describes a user-based usability evaluation of the MITO PACS-viewer and the Kinect used as gesture-input device. This study was performed on surgeons with different specializations and medicine students in a real operating room environment, in which participants had to conduct prototypical tasks on a real patient image study. User satisfaction as well as performance measures such as time-on-task and accuracy were monitored.

5.1 Method

Participants

Eight surgeons and five medicine students (from varying years) from the UMCG participated in this study. The average age of surgeons was 39.3 years old, while students were 22 years old on average. Three of the surgeons were specialized in orthopaedic surgery, two in abdominal surgery, one in thoracic surgery, one in child surgery and one in surgical radiology.

Test environment

The evaluation was conducted in the 'Skills-Centre' at the UMCG, normally used for training purposes (e.g. Crew Resource Management). This room is a good location for a usability evaluation because it resembles the participant's natural working environment and it allows for complete control of external factors that otherwise might have influenced the experiment. The setup of this environment is depicted below in Figure 6.

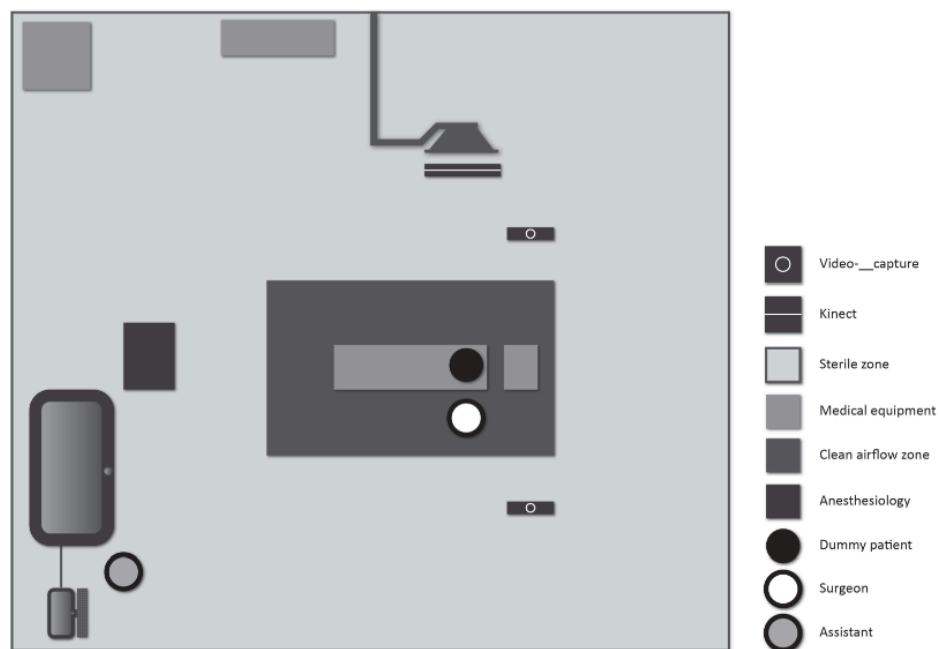


Figure 6. A schematic overview of test environment, depicting the relevant sterile areas, equipment and personnel.

Apparatus

The main computer used for this experiment was a HP Compaq 8200 Elite (Intel Core I5-2400 3.10 GHz, 4GB RAM, Intel HD graphics card) running on Windows 7 Ultimate (64bit) and MITO (version 2.2) connected to a 22 inch NEC MultiSync EA222WMe monitor (1680x1050 resolution), which was placed on a moveable arm already present in the Skills-Centre. The monitor was placed at about head-height so that the surgeon directly faced the setup without having to look up or down. The computer was connected to the Microsoft Kinect for Xbox 360, which was placed at the foot of the monitor and was aimed slightly downwards so that the hands could be tracked while resting next to the participant's body.

Camtasia studio (version 8) was used to record a screen capture of the main computer on which the experiment was conducted, so that exact performance measures could be determined afterwards without having to disturb the experiment. Two Logitech Webcams (Pro 9000) were used to record the participant from the side and from the front so that the participant's actions could be analysed afterwards. Furthermore one laptop was located next to the participant to present the tasks needed to complete the experiment. Finally another laptop, a Dell Latitude E6400 (Intel Core2Duo T9550 2.66 GHz, 4GB RAM, Nvidia Quadro NVS 160m graphics card), was connected to the OR wall-mounted monitor running the same version of MITO, which was used for the control condition.

Materials

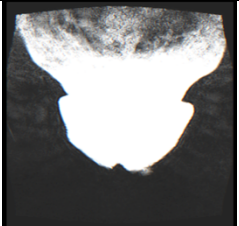
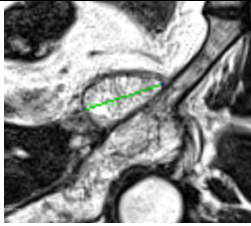
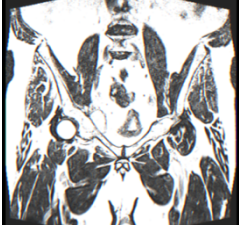
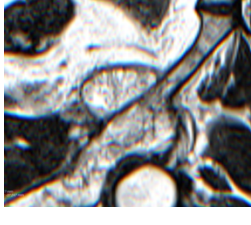

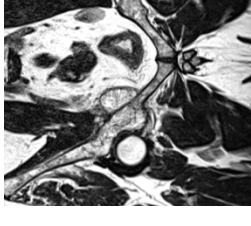
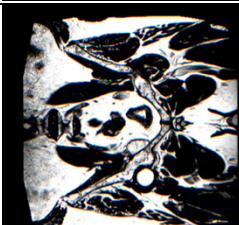
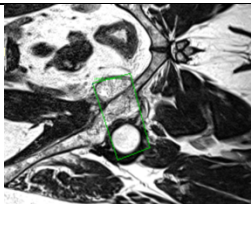
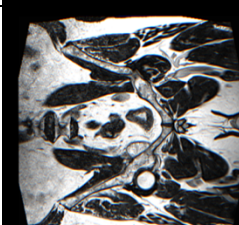
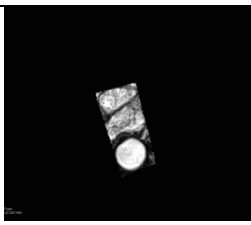

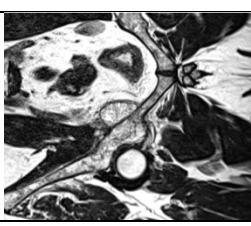
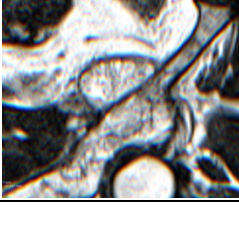
The experiment consisted of three tasks, divided into twelve subtasks, and an extensive questionnaire that had to be filled out. The task set was carefully designed, in consultation with an orthopedic surgeon, to be as representative of a surgeon's work as possible, while taking the limitations of the tested PACS viewer into account. Tasks were also designed to be simple and controlled enough to ensure that all participants are able to understand and perform them.

The main goal for the participant was to locate a tumor in a specific area, to create a clearer view of the tumor and to finally conduct measurements on the tumor (the specific tasks and subtasks are shown in Table 2). Each subtask could be completed with one specific function/gesture. The patient image study used in the tasks was collected from the UMCG database. The images were stored in a local database and all patient information was anonymized accordingly. The study consisted of a MRI scan (using T2 weighting) in the coronal plane of a male patient diagnosed with chondrosarcoma located peri-acetabular.

Before the experiment started, participants were asked about their familiarity with smartphones and tablet/touch computers, their experience with track pad gestures and finally how familiar they were with the Kinect.

After the experiment, participants were asked to fill out an extensive questionnaire, which was designed to measure subjective preferences on the usability of gesture-based interaction in the operating room and consisted of two parts. The first was a system usability scale (SUS) concerning questions about the general usability of the system as a whole. Participants were also asked to rate specific gesture-based interaction properties such as the intuitiveness, physical effort and would-use properties of the used gestures. Participants were asked to indicate the extent to which they agreed with each statement on a seven-point Likert scale. There were also three extra text fields in which participants could indicate positive aspects, negative aspects and additional comments on gesture-based interaction in general (the entire questionnaire can be found in Appendix B.1).

Table 2. The tasks and subtasks used in the experiment. The images and tasks were chosen in consultation with an orthopedic surgeon familiar with the patient study.

Subtask description	Image thumbnail	Subtask description	Image thumbnail
Starting image		3.1. Measure the size of the tumor in the coronal plane as accurately as possible	
1.1. Scroll through the patient study until the following image, containing the tumor, is located		3.2. Clear the selected measurement	
1.2. Orientate the image study to the patient		3.3. Zoom out until the following image is replicated	
2.1. Change the window-level (brightness) until the following image is replicated as accurately as possible		3.4. Select region of interest as depicted in the following image	
2.2. Change the window-width (contrast) until the following image is replicated as accurately as possible		3.5. Cut out region of interest	
2.3. Zoom in on the tumor until the following image is replicated as accurately as possible		3.6. Clear the selected region of interest	
2.4. Translate the tumor to the center of the screen			

Design and procedure

A within subject design was used due to the limited number of participants, with condition and task as within-subject factors. The experiment consisted of two counterbalanced conditions: the gesture-based experimental condition and a control condition in which the participant had to give instructions to an assistant on how to perform the same tasks as in the gesture-based condition. In the control condition the role of assistant was performed by the observer, who was instructed to strictly follow the orders of the participant in order to keep this factor as constant as possible. In the gesture-based condition participants were allowed to ask for the correct gesture if they had repeatedly issued the wrong gesture for the specific task.

One extensively trained user, who repeated the experiment three times, was included in the experiment to obtain an average baseline to which the performance of the surgeons and medicine students could be compared.

First participants were given a short description of the experiment, and were informed that their anonymity was ensured. Before the experiment began, participants were asked to fill out personal information regarding age, handedness and experience with gesture-based interaction and were then given a short training in the gestures needed to accomplish the tasks. This consisted of the experimenter first performing the gestures and demonstrating the system, after which the participant performed all gestures until he or she felt confident using the system. Afterwards participants were asked to calibrate the system. Participants were then shown a laptop aside of them displaying the actual case study and the tasks that needed to be performed. Participants were instructed to perform these tasks as accurately and fast as possible. Each task was divided into several subtasks, which were strictly defined in order to evaluate the performance on certain aspects of the system.

The main goal of task 1 was to locate the tumor and orientate the study to the patient. This task was divided into the following subtasks:

In subtask 1.1, participants had to scroll through the images of the MRI study, until the tumor was located (as shown in the target image).

In subtask 1.2, participants had to orientate the image to the orientation of the patient (90 degrees counterclockwise from the participant's viewpoint).

The main goal of task 2 was to create a clearer view of the tumor. This task was divided into the following subtasks:

In subtask 2.1, participants had to alter the window-level (similar to brightness) until the target image was recreated as accurately as possible.

In subtask 2.2, participants had to alter the window-width (similar to contrast) until the target image was recreated as accurately as possible.

In subtask 2.3, participants had to zoom in on the image until the zoom-level of the target image was recreated as accurately as possible.

In subtask 2.4, participants had to translate the tumor to the center of the viewport as depicted in the target image.

The main goal of task 3 was to measure the tumor and select a region of interest. This task was divided into the following subtasks:

In subtask 3.1, participants had to measure the size of the tumor in the coronal plane (as depicted in the target image).

In sub-task 3.2, participants had to clear the conduct measurement

In subtask 3.3, participants had to zoom out on the image until the target image was replicated.

In subtask 3.4, participants had to select a region of interest, which contained the tumor and the femur-head as depicted in the target image.

In subtask 3.5 participants had to cut out the selected region of interest from the previous subtask.

In subtask 3.6, participants had to finally clear (undo the previous subtask) the selected region of interest.

After the tasks were conducted, participants were asked to fill out the questionnaire and were asked about their experiences with the system, elaborating on possible improvement suggestions.

Data analysis

All data was analyzed with R-Studio (Version 0.97.248). Measured variables included task execution time, task accuracy, gesture recognition accuracy, number of wrongly executed gestures, and the answers to the questions on the questionnaire. Qualitative data was also obtained from the video recordings of the screens and open questions from the questionnaires.

The webcam videos of participants were examined and annotated to determine the start and end times of the defined subtasks. The start time was defined as the moment the participants initiated their movements with their arms into the activation area (see section 4.2 for a detailed explanation). The end time was defined as the moment the last action of the subtask had been performed and the participant started to move their hands out of the activation area.

Accuracy (for task-step 1.1, 1.2, 2.1, 2.2, 2.3, 3.1, 3.3) was determined by calculating the absolute percentage deviation of the acquired value from the target value. Gesture recognition accuracy was determined by observing the video recordings and noting the number of misrecognized (correctly issued) gestures. The same method was applied for determining the number of wrongly issued gestures per task.

Before any statistical analyses between the experimental conditions were performed, first an independent t-test was conducted between the different participants groups (students and surgeons) to see whether the data could be pooled. Afterwards after checking for normality with a Shapiro-Wilk test, a paired t-test was conducted on several performance measures between the experimental and control condition.

To allow for a quantitative analysis of the questionnaire data, the levels of the Likert scale were converted to the following numerical values: strongly disagree = 1, disagree = 2, slightly disagree = 3, neutral = 4, slightly agree = 5, agree = 6 and strongly agree = 7. For each questionnaire the mean response was determined as well as a 95% confidence interval for standard error.

5.2 Performance measure results

In this section an overview of the performance data is presented, such as task execution time, accuracy and amount of wrongly recognized gestures. Before the results of both type participants (surgeons and students) were combined, independent t-test were conducted to see whether there were any significant performance differences between these two groups. A two-way Anova with Completion Time as dependent variable, with Condition and Group as main effects revealed no significant effect for Group ($F_1 = 2.02$, $p > 0.05$), and no interaction effect between Condition*Group ($F_1 = 0.06$, $p > 0.05$). This also holds true for Accuracy, which revealed no significant effect for Group ($F_1 = 1.20$, $p > 0.05$), and no interaction effect between Condition*Group ($F_1 = 0.12$, $p > 0.05$). These results indicate that data from both groups can be pooled together, and will be presented as such below.

Furthermore to check whether the experiment was indeed counterbalanced correctly an analysis of variance was also conducted between the group that started in the gesture condition ($N=7$) and the group that ended in the gesture condition ($N=6$). A two-way Anova with Completion Time as dependent variable and Condition and Order as main effects revealed no significant effect for Order ($F_1 = 2.53$, $p > 0.05$), and no interaction effect between Condition*Order ($F_1 = 0.00$, $p > 0.05$). For Accuracy no significant effect was found for Order ($F_1 = 3.78$, $p > 0.05$), but there was a significant interaction effect between Condition*Order ($F_1 = 6.68$, $p = 0.02$). Participants who conducted the gesture condition after the control condition were more accurate than participants who conducted the gesture condition before the control condition. This significant effect in the gesture condition might be due to a slightly unbalanced design or due to some possible outliers in the accuracy data of participants who conducted the gesture condition first. Because no explanation could be found and because the task time showed no abnormalities, the interaction effect in the accuracy data will be attributed to chance and will not be corrected for.

Task execution time

When we look at the average completion time per task-step/action as shown in Figure 7., it indicates that the largest differences between the two conditions can be found in the measurement and ROI selection tasks, and to a lesser extent in the animate and translate tasks. Some tasks are slightly faster in the gesture condition on average compared to the control condition, such as for changing window-width, rotating and zooming in.

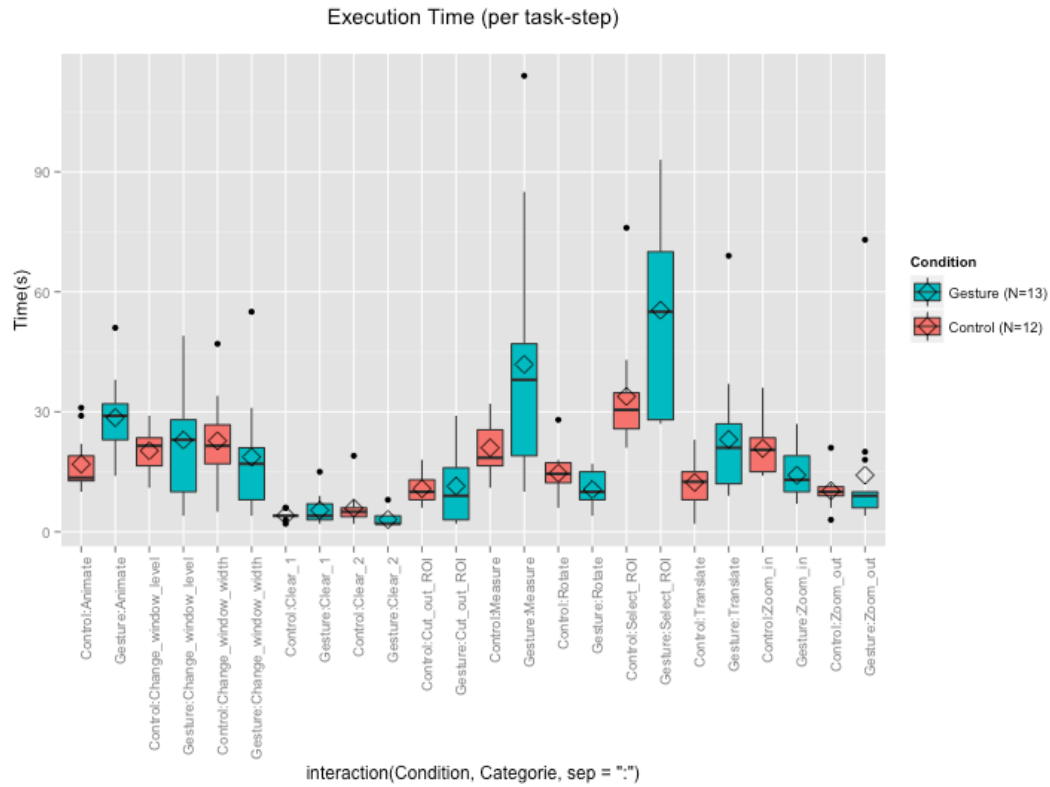


Figure 7. Boxplots depicting task-step/action completion time in both the control condition and the gesture condition. The centre line depicts the median whereas the diamond depicts the mean value. Error bars indicate standard errors.

Figure 8. shows that participants are faster in the control condition than in the gesture condition for total completion time. The total completion time was calculated by adding up all separate task-step execution times. A Shapiro-Wilk normality test indicated that the observed values were normally distributed ($W = 0.9301$, $p > 0.05$), after which a paired t-test was performed, showing a significant difference in average completion time between the control and experimental condition ($t_{11} = 2.94$, $p = 0.01$) in which the average control condition (195.9s) was 1.25 times faster than the average gesture condition (245.8s).

Figure 8. also depicts a baseline in which an extensively trained participant conducted the same tasks three times in the gesture condition. This baseline indicates the possible effect of training on completion time and Figure 8. shows that this baseline is faster than both control and gesture condition by novice participants. A one-sample t-test of the (novice) gesture condition compared to the baseline (113.7 seconds) show a significant effect ($t_{12} = 11.34$, $p < 0.001$). This also holds true for the (novice) control condition compared to the same baseline ($t_{11} = 6.69$, $p < 0.001$). The baseline (114s) is 1.72 times faster than the control condition (195.9s) and 2.18 times faster than the gesture-based condition.

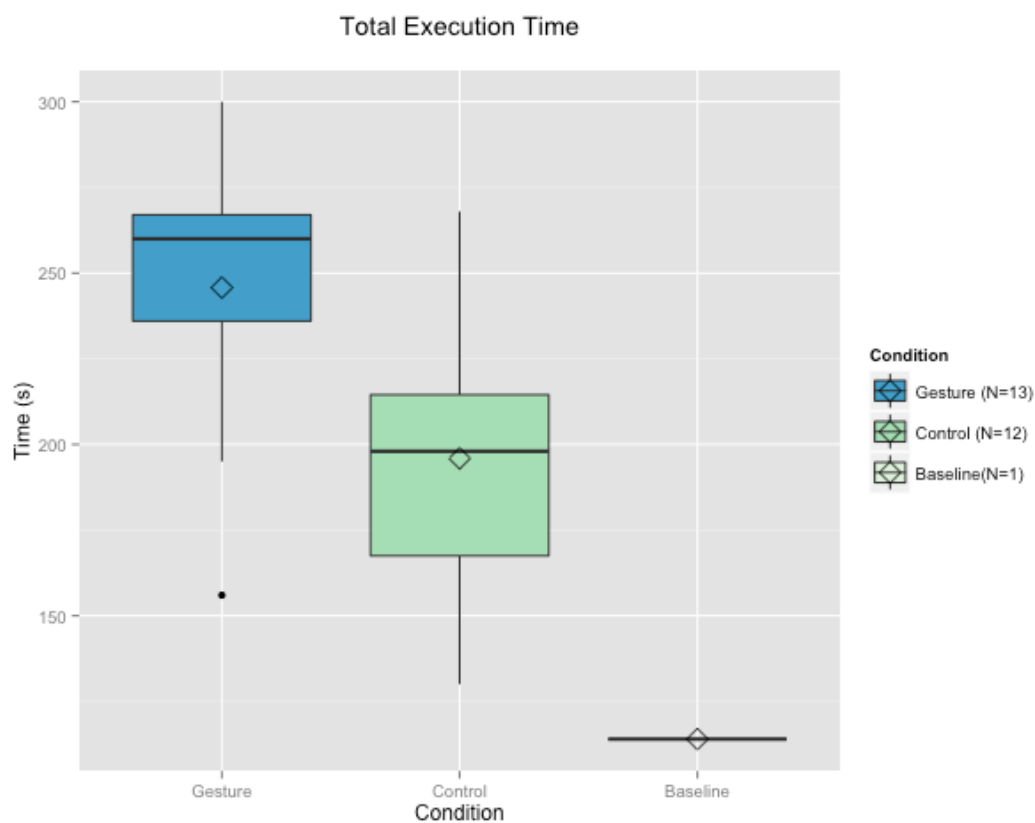


Figure 8. Boxplots depicting the total completion time in both the control condition and the gesture condition. The centre line depicts the median whereas the diamond depicts the mean value. Error bars indicate standard errors. NB: A baseline condition is included (1 extensively trained participant on 3 trials) as a reference point.

Accuracy

Figure 9. shows the accuracy of relevant task-steps (see Table 3 tasks 1, 2, 3, 4, 5, 7, 9) grouped by condition (a baseline condition of an extensively trained participant is also included). The figure shows that changing window-width and window-level, measuring and zooming-out in the gesture condition are less accurate compared to the control condition. Although the baseline indicates that the accuracy of the control condition or even more accurate results (for zooming) can be obtained for all task-steps in the gesture condition. Furthermore the animating, rotating and zooming task-steps show similar results between the gesture and control condition.

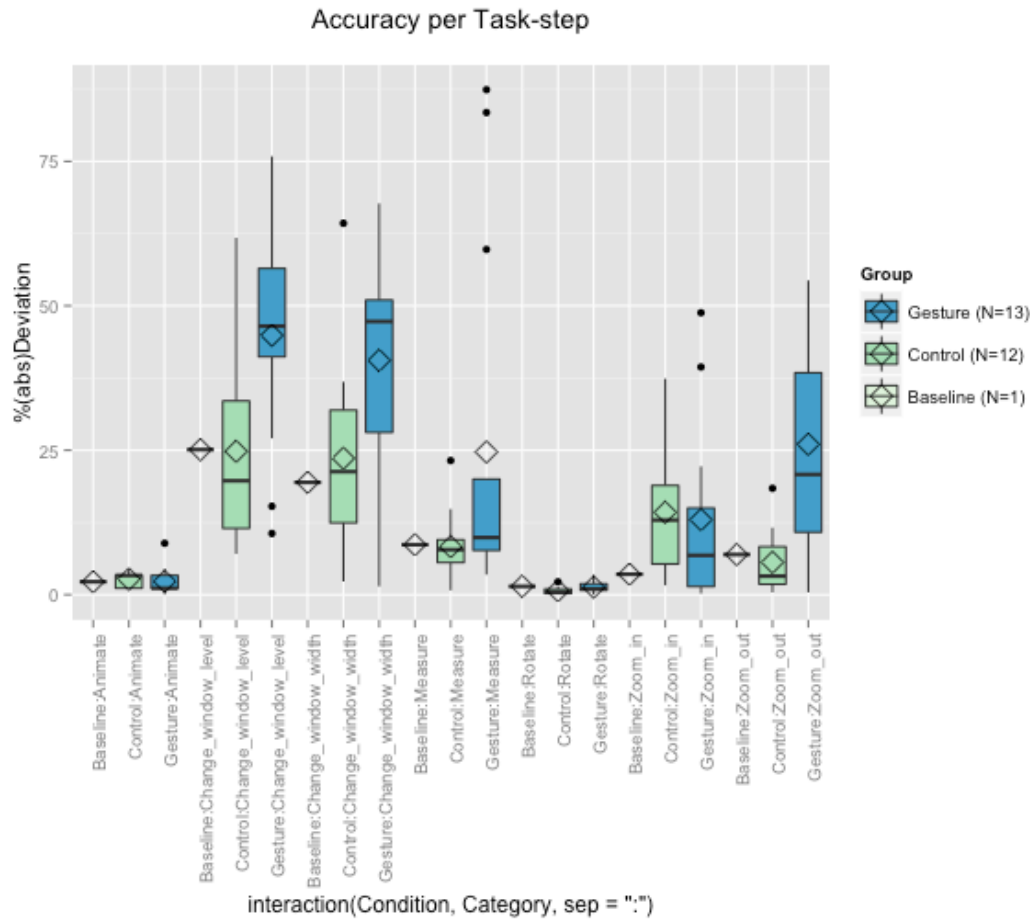


Figure 9. Boxplots depicting the accuracy (% absolute deviation) of relevant task-steps in both the control condition and the gesture condition. The centre line depicts the median whereas the diamond depicts the mean value. Error bars indicate standard errors. NB: A baseline condition is included (1 extensively trained participant on 3 trials) as a reference point.

Number of misrecognized gestures

To see how many gestures are misrecognized by the system, Figure 10. shows the amount of misrecognized gestures per task-step in the gesture condition (a baseline condition of an extensively trained participant is also included). The figure shows that the number of misrecognized gestures is highest for respectively selecting a region of interest, conducting measurements followed by cutting out the region of interest. These observations are in line with the values obtained in the baseline condition.

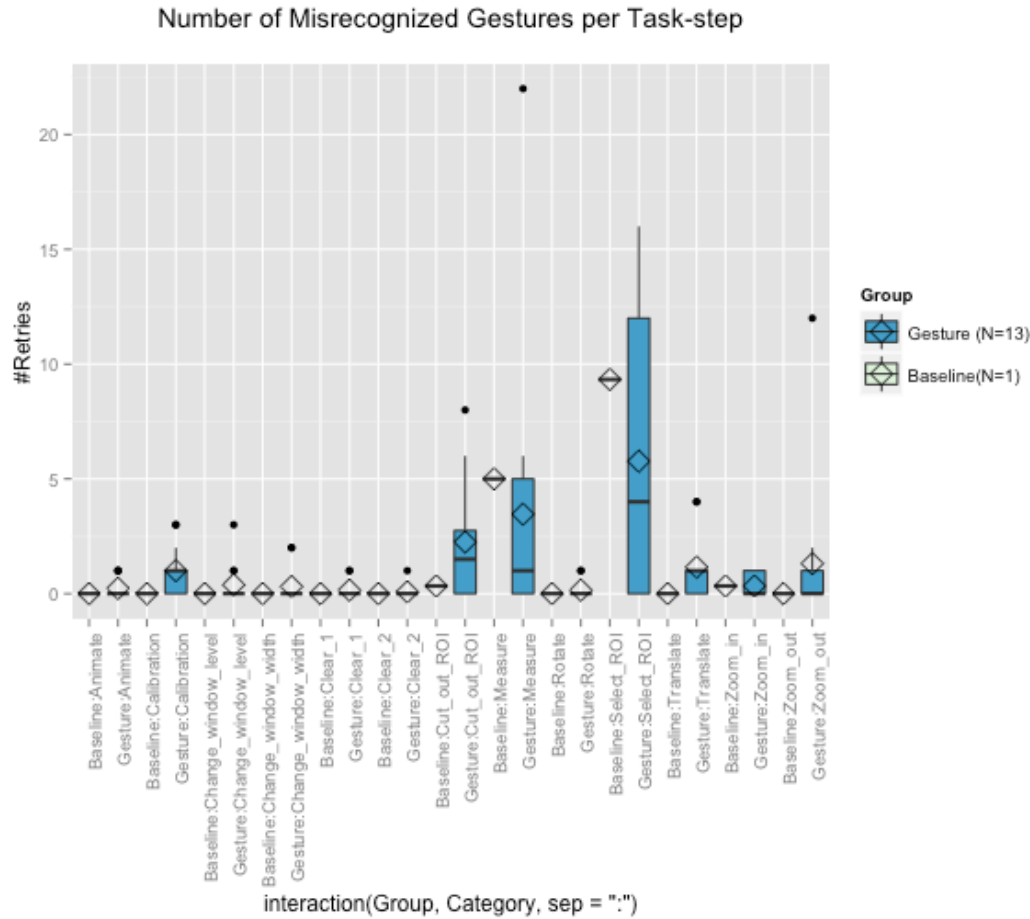


Figure 10. Boxplots depicting the number of misrecognized gestures by the tested system in the gesture condition. The centre line depicts the median whereas the diamond depicts the mean value. Error bars indicate standard errors. NB: A baseline condition is included (1 extensively trained participant on 3 trials) as a reference point.

Number of incorrectly issued gestures

In order to see how many gestures are incorrectly issued by the participant can be found in Figure 11. This figure shows the average amount of wrong gestures issued by the participant per task. Visual inspection of the figure indicates that this number is highest for animating through the patient study and clearing selections. The baseline condition is not taken into account in this figure because no gestures were issued wrong, as can be expected from an extensively trained participant.

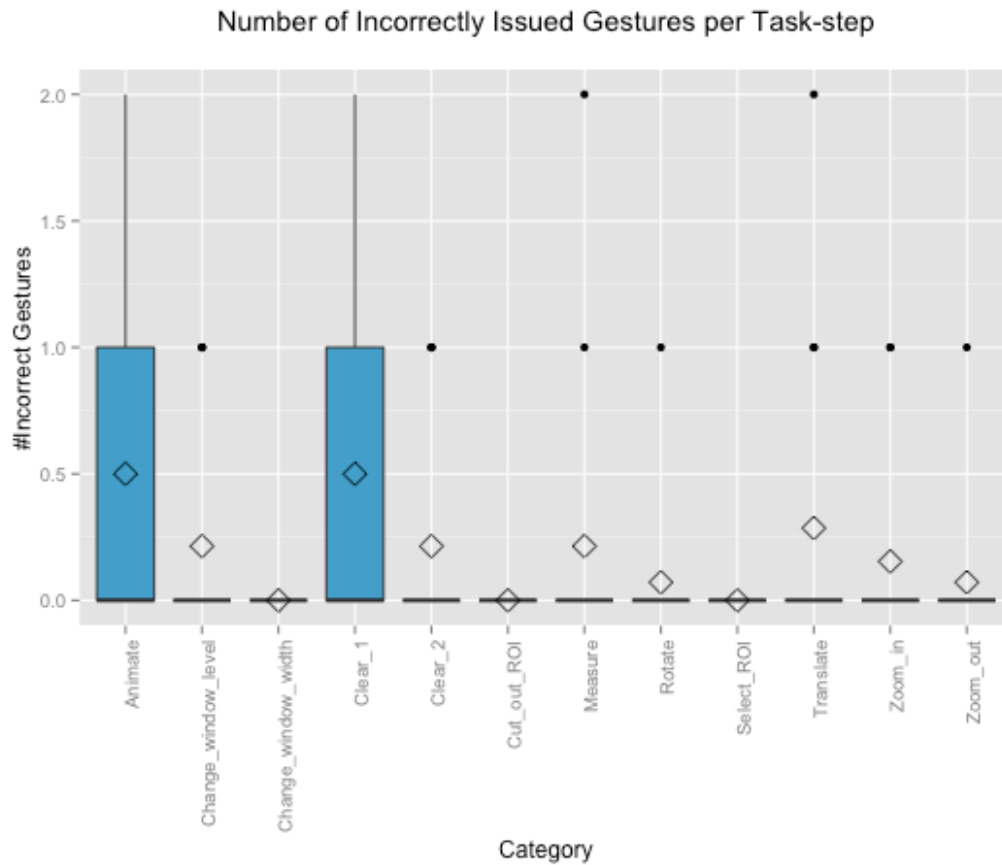


Figure 11. Boxplots depicting the number of incorrect gestures issued by the participant in the gesture condition per task. The centre line depicts the median whereas the diamond depicts the mean value. Error bars indicate standard errors.

5.3 Usability results

In this section an overview of acquired usability measurements is presented, including general system usability questions as more specific questions concerning the intuitiveness of the used gestures and open questions regarding possible improvements for future systems.

Pre-experimental familiarity with gesture-based interaction

Figure 12. gives an overview of the participants' familiarity with gesture-based interaction. Participants were asked how familiar they were (ranging from 1 = not familiar to 7 = very familiar) with smartphones, tablet and touch computers, track pad gesture experience and finally how familiar they were with the Kinect already. Figure 11 indicates that the participants were on average very experienced with device-based gesture technology whereas participants were hardly familiar with deviceless gesture-based technology such as the Kinect.

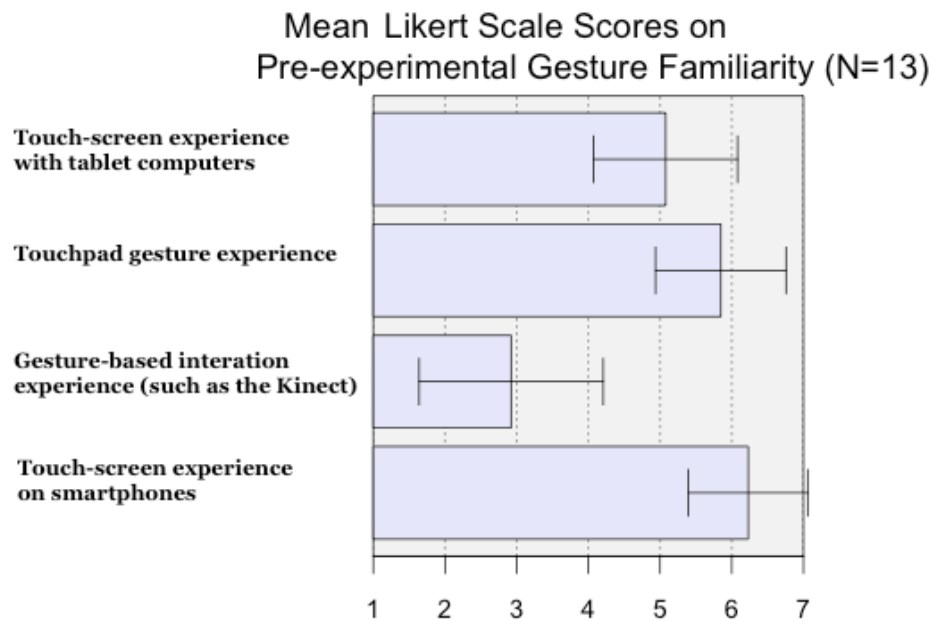


Figure 12. Bar chart depicting the participants' familiarity with gesture-based interaction before the experiment. The centre line depicts the mean value and error bars indicate a 95% confidence interval of the standard error.

System Usability Scale (SUS) response

After the experiment was conducted participants were asked to fill out a System Usability Scale (SUS), which is a popular scale often used in usability research. This questionnaire deals with ten alternating negative and positive formulated statements such as “I had to learn a lot before I could use this system” and “I thought the system was easy to use”. The results have been recoded for positive outcome, and can be found in Figure 13.

Responses in general show a positive evaluation of the system’s usability, with all average Likert scores greater than 4. The least positive scores (ranging from indifferent to moderately positive) were given to the statements concerning if the functions were well implemented and the confidence of the participants while using the system.

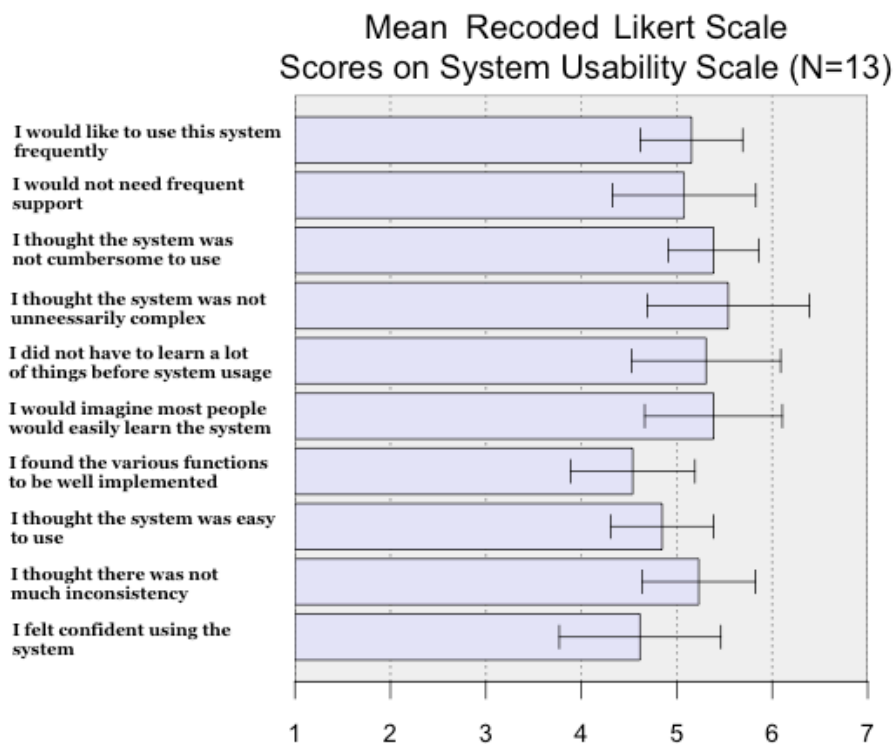


Figure 13. Bar chart depicting the Likert scale responses on statements used in the System Usability Scale (SUS). The centre line depicts the mean value and error bars indicate a 95% confidence interval of the standard error.

Gesture specific questionnaire response

Participants were asked to fill out a questionnaire concerning the usability of specific gestures. Participants were given statements whether they the relevant gesture was intuitive, if the gesture required much physical effort and whether they would want to use this gesture in this form in a real operating room situation. The results have been recoded for positive outcome and can be found in Figure 14.

In general responses show a positive evaluation of the implemented gestures. Nearly all gestures were rated positively for their intuitiveness, physical effort and would-use. Especially zooming, rotation and clearing a selection were rated highest on all three properties, followed by pointing, animating and translation. The least positive score (ranging from indifferent to moderately positive) were the gestures for changing window-width and window-level, cutting out a region of interest and finally selection, which scored moderately positive on intuitiveness and would-use but moderately negative on physical ease.

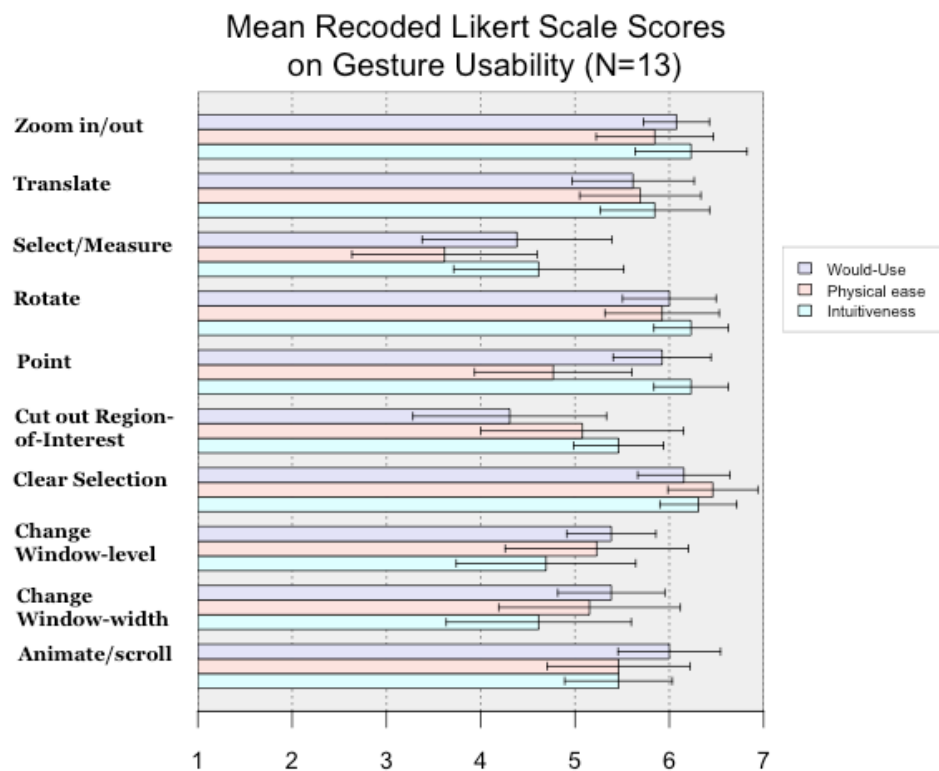


Figure 14. Bar chart depicting the Likert scale responses on three properties (would-use, physical ease and intuitiveness) concerning the usability of the used gestures. The centre line depicts the mean value and error bars indicate a 95% confidence interval of the standard error.

Gesture-based interaction questionnaire response

Participants were also asked to fill out a questionnaire concerning system specific questions similar to the above-described SUS. This questionnaire deals with eight alternating negative and positive formulated statements such as “System feedback was not sufficient” and “(De)activation of gestures was well implemented”. The results have been corrected for positive outcome, and can be found in Figure 15.

Again in general this figure shows a positive evaluation of the gesture-based specific properties of the tested system. Participants responded very positive when asked whether they believed that gesture-based interaction could eventually be implemented in the real OR and if they already preferred the current system to asking an assistant (as in the control condition). Also positive were the reactions to the statements concerning sufficient system feedback, lack of physical feedback and whether the participant considered the used gestures to be applicable in a real OR setting. Participants were least positive about the way the calibration of the tested system was implemented.

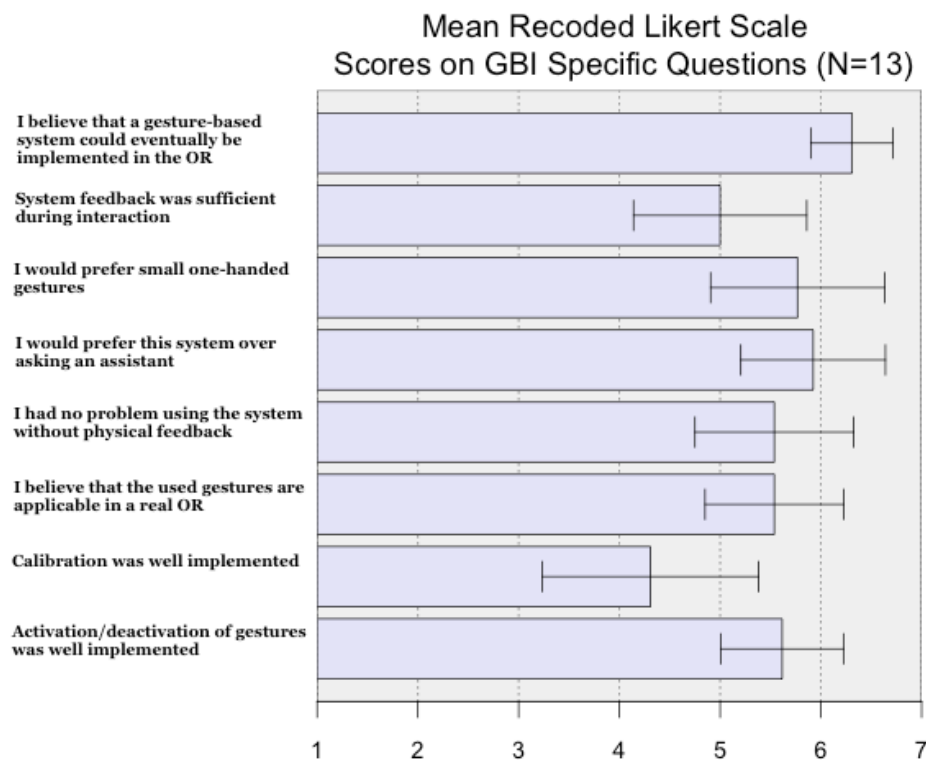


Figure 15. Bar chart depicting the Likert scale responses on system specific usability statements. The centre line depicts the mean value and error bars indicate a 95% confidence interval of the standard error.

Open question responses

The questionnaire was concluded with four open questions in which participants were asked to fill out any functionality that was missing in the tested system. Also they were asked what they regarded as positive as well as negative aspects of gesture-based interaction in general. Finally there was room for remaining remarks or suggestions for future developments. The results of the open questions are presented below. The number between brackets indicates how often the particular answer was given.

Table 3. Answers on the first open question regarding missing functionality, divided on participant type (surgeon or student).

Did you miss any functionality?	
Surgeons (N=8)	Students (N=5)
Switching between study or image source type (x5)	Switching between study or image source type
Pause interaction (x3)	Full mouse control (left and right click functionality)
Multiple viewports (2x)	
Motion picture support (angiographic studies for example)	
Placing two studies next to each other (old/new)	
3D image reconstruction	
Light control would be useful	
Patient information on screen	
Accessing other relevant patient information	

Table 4. Answers on the second open question regarding positive aspects of gesture-based interaction for operating room purposes, divided on participant type (surgeon or student).

What are positive aspects of gesture-based interaction in the OR?	
Surgeons (N=8)	Students (N=5)
Direct control (x3)	Direct control (x2)
Independence (x3)	Independence (x2)
Sterile (x3)	Sterile
Intuitive (x2)	Intuitive
Fast (x2)	Fast
Wireless	Easy
Deviceless technology	

Table 5. Answers on the third open question regarding negative aspects of gesture-based interaction for operating room purposes, divided on participant type (surgeon or student).

What are negative aspects of gesture-based interaction in the OR?	
Surgeons (N=8)	Students (N=5)
Less accurate (x3)	Less accurate (x3)
Large movements may compromise sterility (x2)	
Risk of sterility breach (x2)	
System requires a lot of space in small OR environment	System requires a lot of space in small OR environment
Steep learning curve	
Calibration requires a lot of effort	Calibration requires a lot of effort
What do to when system crashes?	

Table 6. Answers on the fourth and last open question regarding remaining remarks or suggestions for gesture-based interaction in the operating room, divided on participant type (surgeon or student).

Do you have any other remaining remarks or suggestions?	
Surgeons	Students
For future purposes, integration with intra-operative navigation would be very useful	Window-level and window-width gestures should be more similar, they are counter-intuitive compared to each other
Selecting ROI in this way is not useful	Transition between functionality without activation/deactivation of the current gesture
Better system feedback	Better system feedback
Scrolling in different directions with one open palm through the body axis would be more intuitive	
Movements preferred below chest area (due to sterility)	
A lot of functions are hardly used (measuring is often done before surgery), this system is especially useful for pointing out an area of interest to colleagues etc.	
System that recognized gestures from below would be preferred	
To deactivate the system, keeping hands at certain position for 2 seconds would be useful	
A good sterilizable mouse would be just as useful in the OR	
A better gesture-detection algorithm would be preferred over the current (de)activation method	
Picture rotates over image axis and not over view axis, this would be more intuitive while zoomed in for example	

5.4 Summary

The most important results of the usability evaluation can be summarized in the following three groups:

Performance measure results

The task completion time results clearly show that the current control condition is significantly faster than using this particular gesture-based system. Striking however is that an extensively trained participant is significantly faster with the gesture system when compared to novice participants but also compared to the control condition. Furthermore accuracy measures show that the tested system is less accurate than the control condition on several aspects, especially changing window-level, window-width and conducting measurements are less accurate. Again baseline accuracy values are visibly higher than the accuracy values acquired by the novice participants. The amount of incorrectly issued gestures by the participants is highest for animating and clearing a selection. And the amount of incorrectly recognized gestures by the system is highest for selection and clearing a selection.

Usability results

The several different usability questionnaires show that participants are generally positive about the tested gesture-based system. The System Usability Scale (SUS) was

generally filled out positive, although participants responded slightly less positive about their confidence while using the system and the implementation of the functionality could be better. The results of the questionnaire concerning the usability of the implemented gestures were also generally positive, although the gestures for changing window-level, window-width, selection and cutting out a region of interest could be improved. Finally participants responded generally positive to questions concerning gesture-based system specific questions. They responded positive to statements such as whether they already preferred the current system to asking an assistant and if they believed that gesture-based interaction would eventually be implemented in the real OR (feasibility of GBI in the OR). Questions regarding physical feedback and system feedback were also answered positively. Interestingly participants stated that they would prefer smaller one-handed gestures and the calibration step of this system could be improved.

Open question results

From the open questions it is clear that the functionality of the tested gesture-based system should be extended with the possibility to switch between studies and image source type, to display two patient studies next to each other and to allow for multiple viewports. Furthermore participants would like to see a pause button, in order to turn interaction off and on. Other interesting suggestions included the integration of gesture-based light control and the ability to bring up patient information. Further suggestions for improvement of a gesture-based system include adjusting the scroll gesture to allow for moving the hand in the direction the images were taken in. Also participants would rather use gesturing below chest height, which would be less tiring and more importantly less prone to contamination of the sterile working environment, this also holds true for the implemented ROI selection gesture which might touch the chest area. Finally when participants were asked to identify the positive aspects of gesture-based interaction, their responses corresponded to the hypothesized advantages of such systems. Participants thought that gesture-based interaction offered them direct control, independence and a sterile, intuitive way of interacting with the patient image studies. Negative aspects of gesture-based interaction, or at least this gesture-based system, are its decreased accuracy, steeper learning curve and the chance of compromising sterility when too large gestures are used.

Due to the inaccuracy of selection (measurements) in the current system it is important to take a closer look at this aspect of gesture-based interaction. This is very important because users would like to be able to correctly and accurately make selections, and in the case of operating room purposes be able to conduct accurate measurements. To this extent a second study was conducted described in the next chapter, which looks at different popular selection techniques using the Kinect with different settings in order to get a better, and more quantified overview of the possibilities and limitations of the Kinect on it's accuracy properties.

Chapter 6.

Performance evaluation of gesture-based selection techniques

Because of the importance of selection in user interfaces and the particularly challenging nature of selection in gesture-based interaction, due to the absence of a physical selection mechanism and absence of haptic feedback, it is important to be able to quantify and compare different selection techniques on several performance measures and usability properties. This chapter deals with such a performance evaluation of several gesture-based selection techniques using the Kinect as input device. This experiment was conducted in line with the ISO 9241 – part 9 standard, a standard that establishes uniform guidelines and testing procedures for evaluating computer pointing devices (see section 2.3 for a detailed description). In this experiment two popular Kinect selection techniques, ‘Dwell’ and ‘Push’, are compared to using the mouse as selection tool on accuracy, error and usability measures. Furthermore because in surgeon-computer interaction the movement area is constrained, different sized interaction areas are compared to see how this influences the performance measures.

In this experiment, participants were asked to perform a multi-directional point and select task, inspired by the task used in a study on “Evaluating Measures for Evaluating Computer Pointing Devices” by Mackenzie, Kauppinen and Silfverberg (2001). In their study participants had to conduct such a multi-directional selection task in which four different pointing devices were evaluated on movement time, error rate, throughput and seven other performance measures of interest for capturing movement behaviour during trials. In the present study participants had to conduct a similar experiment in which they moved the cursor with their dominant hand and selected one of the highlighted targets with their non-dominant hand. In this case we are solely interested in measuring movement time, error rate and throughput due to their priority during surgeon-computer interaction.

6.1 Method

Participants

10 volunteer students (8 male, 2 female) were recruited from the University of Groningen. Participants ranged from 21 to 28 years (mean = 24.3). All were daily users of computers, reporting 6 to 9 hours usage per day (mean = 7.1). Nearly all participants were familiar with the Kinect and had varying experience using it, ranging from 1 to 7 on a seven-point Likert (1 = not familiar, 7 = very familiar, mean = 4.1).

Test environment

The experiment was conducted in a suitable office in the UMCG, in which the participant could not be disturbed.

Apparatus

The main computer used for this experiment was a HP Compaq 8200 Elite (Intel Core I5-2400 3.10 GHz, 4GB RAM, Intel HD Graphics card) running on Windows 7 Ultimate (64bit) connected to a 19 inch Philips Brilliance 170P monitor (1280x1024 resolution). Furthermore the Kinect for Windows was used, which was used as input for the Alces Universal Gesture Mouse software¹. This program was used to control the Windows cursor in the multi-directional selection task. The program functions as a plugin, translating movements of the user into cursor movements and click events. Several settings make this program suitable for this study, such as setting the wanted selection technique (in this case ‘Dwell’ and ‘Push’) and changing the area of interaction to different pre-sets (in this case 2x and 4x, see Figure 15). Finally the multi-directional

¹ <http://www.alcestech.com/universal-gesture-mouse.html>

selection task itself was implemented in the C# programming language using the Microsoft Visual Studio 2012 Professional IDE (64bit). See Appendix D.1 for a detailed description of the code.

The two gesture-based selection techniques that were evaluated in this experiment were the ‘Dwell’ and ‘Push’ technique. Selection in the Push condition occurs when the participants pushes his or her non-dominant hand towards the screen. Feedback of this process can be seen on-screen, due to a red filled circle around the cursor (see Figure 16), which becomes smaller as the participant’s hand is pushed forward until a threshold is reached and the inner filling disappears. Selection in the Dwell condition occurs when the participants holds his or her hand still for 1.5 seconds. Feedback of this process can also be seen on the screen; this time the red-filled circle surrounding the cursor starts filling up green bottom-up wise until it is completely filled green, indicating a selection.

Materials

Participants were asked to fill out a questionnaire after completing each condition concerning the usability of the interaction technique. Questions included the perceived accuracy of the technique, if arm fatigue was high during interaction and if they would like to use this interaction technique for example (the complete questionnaire can be found in Appendix D.2. Participants could answer these statements on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree).

Design and procedure

The experiment was a 5 x 3 x 10 within-subjects factorial design. Consisting of 5 conditions (Dwell x2, Dwell x4, Push x2, Push x4, Mouse), 3 sequences of 16 trials and 10 participants per experiment. The total number of trials amounted to 2250 (10 participants x 5 conditions x 3 sequences x 15 trials).

In this experiment a distinction was made between three different selection methods and two different interaction areas. The selection methods consisted of two gesture-techniques ‘Dwell’ and ‘Push’ and one condition in which the participant used a mouse to select the targets. The two interaction conditions differed in the area in which the participant interacts with the cursor, namely a ‘2x’ and a ‘4x’ interaction area as can be seen in Figure 15: the 4x condition had a four times larger interaction area compared to the 2x condition.

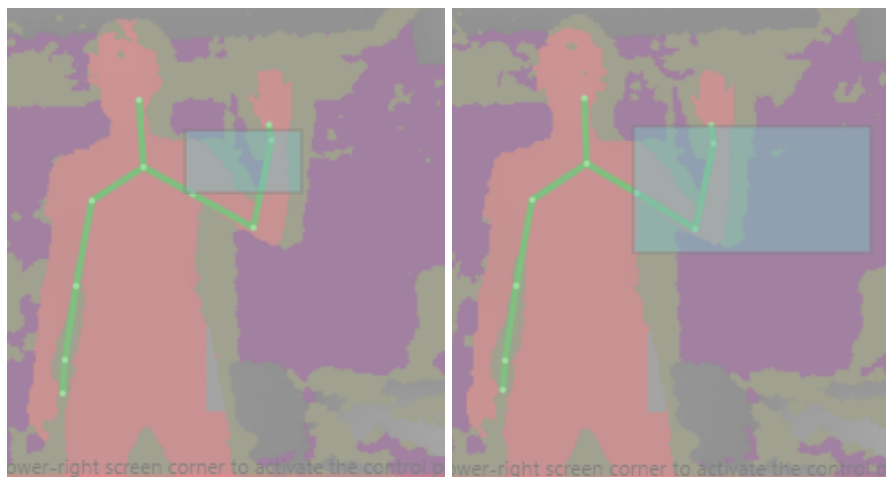


Figure 16. A participant controlling the cursor in the ‘2x’ and ‘4x’ interaction area condition. The 4x condition (right) entails a four times larger interaction area compared to the 2x condition (left).

Each participant was tested with all different methods. The order of conditions differed for each participant according to a balanced Latin square design. Before the experiment began, participants were given a short description of the experiment and were asked to

fill out a questionnaire concerning their age and familiarity with the Kinect and were given a short training in the gestures needed to accomplish the tasks. Finally participants were asked to stand 2 meters from the screen, indicated by a marker, in order to resemble an average operating room range.

The task was the multidirectional point-select task described in ISO 9241-9 (Iso, 1998)(see Figure 17.). This task involves 16 circular targets, of which the entire circle on which the targets were located had a diameter of 660 pixels, whereas each target circle had a diameter of 40 pixels. Because the goal of this experiment was to test the accuracy of several selection techniques, only one task condition was used with a nominal difficulty index of 4.13 bits.

The task for the participant was to select the highlighted target circle. A trial began after the top target was selected. The next target was the circle on the opposite side of the previous target plus one circle clockwise, and so forth. Every new target was colored red, which changed from target to target as the sequence progressed until the final target was reached at the top and was colored blue. At this point participants could take a rest, due to the physically challenging nature of the task. Participants were instructed to select the targets as quickly and accurately as possible. Each participant was tested for all selection methods on three sequences of 15 target selections. Finally after each condition participants were asked to fill out a questionnaire concerning the usability of the interaction technique.

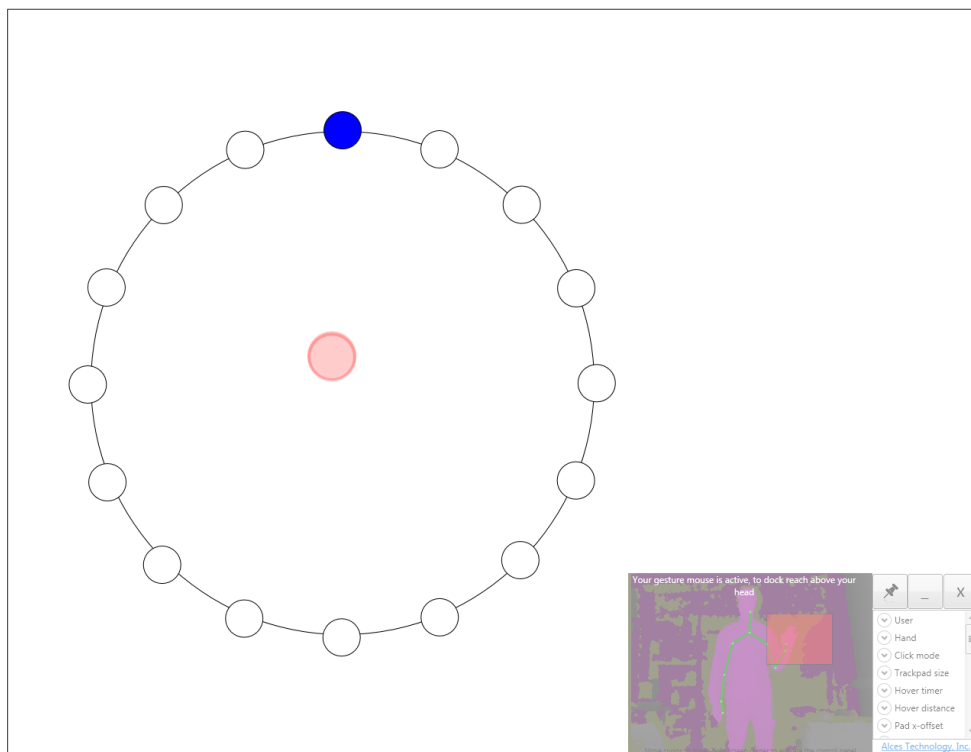


Figure 17. The experiment task with (on the left) the circular layout consisting of 16 selection targets and (on the right) a feedback screen showing the tracked user and the gesture interaction zone (coloured red). In the centre of the circular layout the red filled circle resembles the hand, which the cursor follows and also acts as a visual feedback mechanism.

Data analysis

All data was analyzed with R-Studio (Version 0.97.248). Measured variables included accuracy, movement time (MT), throughput, number of errors and the answers to the questions on the questionnaire.

Movement time (also referred to as selection time) was defined as the time needed between two subsequent selections of target circles. It is important to note that data collection began after the first selection and thus not for the top target (which was used as start and stop target between trials). Accuracy was determined by calculating the number of pixels that the click-event was registered from the exact center of the target circle. And the number of errors was determined by counting how many click-events took place outside the border of the target circle. According to the ISO 9241-9 standard only throughput is proposed as performance measure. Throughput is a combined measure derived from speed and accuracy and is defined as follows:

- $Throughput = ID_e / MT$
- $ID_e = \log_2(D/W_e + 1)$.
- $W_e = 4.1333 \times SD_x$

In which ID_e is the effective index of difficulty in bits, which in this case was 4.13 bits. D refers to the distance to the target and W_e refers to the effective width of the target. The effective width is actually calculated from the distribution of selection coordinates over a sequence of trials. This is reflected in SD_x , which is the standard deviation in selection coordinates measured along the axis of target approach. This indirectly implies that W_e reflects the spatial variability or accuracy, and thus throughput resembles a measure of both speed and accuracy (MacKenzie et al., 2001).

Because the results deal with reaction time data, which is often skewed (which was also the case for the collected data), outliers were removed with an R-script. This script removed data entries that deviated more than three standard deviations from the mean. A standard deviation cutoff is advised when subject variability is large (Ratcliff, 1993).

All data was checked for normality with a Shapiro-Wilk test, and was subjected to several two-way analyses of variance between conditions, with subjects as blocking factor due to the repeated measures design.

To allow for a quantitative analysis of the questionnaire data, the levels of the Likert scale were converted to numerical values as follows: strongly disagree = 1, disagree = 2, slightly disagree = 3, neutral = 4, slightly agree = 5, agree = 6 and strongly agree = 7. For each questionnaire the mean response was determined as well as a 95% confidence interval for standard error.

6.2 Performance measure results

In this section an overview of the performance data is presented, such as movement time, accuracy, number of errors and usability. Before the results were subjected to statistical tests, it was first checked for normality. After removing outliers (3x SD) the following results on the Shapiro-Wilk tests were obtained:

For selection time: Dwell-2x ($W = 0.96, p > 0.05$), Dwell-4x ($W = 0.91, p > 0.05$), Push-2x ($W = 0.90, p > 0.05$), Push-4x ($W = 0.93, p > 0.05$) and Mouse ($W = 0.86, p > 0.05$).

For accuracy: Dwell-2x ($W = 0.93, p > 0.05$), Dwell-4x ($W = 0.97, p > 0.05$), Push-2x ($W = 0.92, p > 0.05$), Push-4x ($W = 0.94, p > 0.05$) and Mouse ($W = 0.89, p > 0.05$).

For error: Dwell-2x ($W = 0.87, p > 0.05$), Dwell-4x ($W = 0.89, p > 0.05$), Push-2x ($W = 0.85, p > 0.05$), Push-4x ($W = 0.93, p > 0.05$) and Mouse ($W = 0.54, p < 0.001$).

For throughput: Dwell-2x ($W = 0.94, p > 0.05$), Dwell-4x ($W = 0.92, p > 0.05$), Push-2x ($W = 0.98, p > 0.05$), Push-4x ($W = 0.94, p > 0.05$) and Mouse ($W = 0.84, p < 0.001$).

These results show that nearly all data is normally distributed except for error and throughput in the mouse condition. The error can be explained by the superior accuracy of the mouse, nearly all participants conducted the tasks flawless, but two participants had slightly higher error scores, which probably caused departure from normality. The slight departure from normality for throughput ($W = 0.84$) can probably be explained because it is the product of selection time ($W = 0.86$) and accuracy ($W = 0.89$). These results generally show no normality concerns, so the data below will be presented as such.

Selection time

Figure 18. shows the average selection time per condition per sequence. A repeated measures analysis of variance, with condition as main effect and subject as blocking factor, indicated a significant effect between conditions ($F_4 = 16.18$, $p < 0.001$). To find out which condition(s) differed significantly, linear contrasts were formed, after which an F ratio was obtained, corrected for Type-I errors and finally a Tukey probability was calculated. The following results were obtained (Table 7):

Table 7. Results of the post-hoc tests on selection time, obtained by forming linear contrasts, computing an F ratio, correcting for Type-I errors and finally computing a Tukey probability. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

F-ratio and significance codes	Mouse	Dwell-2x	Dwell-4x	Push-2x	Push-4x
Mouse	-	207.3***	89.7***	165.7***	99.5***
Dwell-2x	207.3***	-	24.3***	2.3	19.6***
Dwell-4x	89.7***	24.3***	-	11.6*	0.3
Push-2x	165.7***	2.3	11.6*	-	8.4*
Push-4x	99.5***	19.6***	0.3	8.4*	-

These results indicate that there is a significant difference between nearly each condition, except between Dwell-2x vs. Push 2x and Dwell-4x vs. Push 4x. The mouse is the fastest interaction technique (1060ms), followed respectively by the Dwell-4x (3173ms) and Push-4x (3329ms) interaction technique and finally by Push-2x (3992ms) and Dwell-2x (4290ms). Furthermore both '4x' conditions show significantly faster selection times when compared to both '2x' conditions, while participants have to interact with their hands in a larger tracking area. A repeated measures analysis of variance on selection time, with trial and condition as main effects and subject as blocking factor revealed a significant effect ($F_1 = 11.39$, $p < 0.001$), indicating a learning effect in which participants become faster over time in these three trials (1st trial: 3621ms, 2nd trial: 3450ms, 3rd trial: 3377ms on average over all conditions).

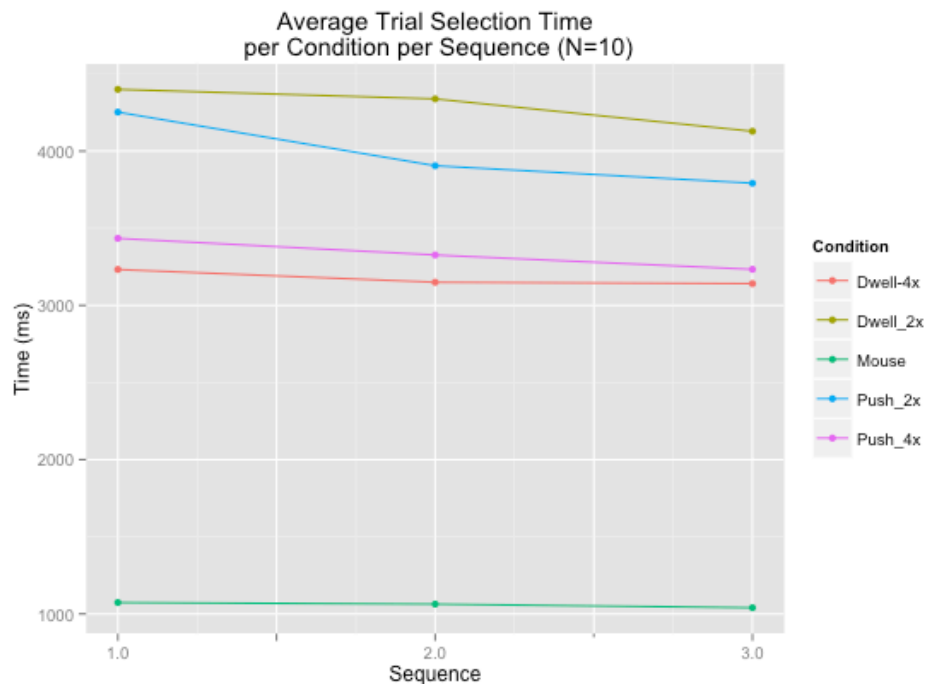


Figure 18. Average trial selection time per condition over three different sequences. The y-axis indicates the average time between each selection in milliseconds.

Accuracy

Figure 19. shows the average selection time per condition per sequence. A repeated measures analysis of variance, with condition as main effect and subject as blocking factor, indicated a significant effect between conditions ($F_4 = 11.87$, $p < 0.001$). To find out which condition(s) differed significantly, linear contrasts were formed, after which an F ratio was obtained, corrected for Type-I errors and finally a Tukey probability was calculated. The following results were obtained (Table 9):

Table 8. Results of the post-hoc tests on accuracy, obtained by forming linear contrasts, computing an F ratio, correcting for Type-I errors and finally computing a Tukey probability. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

F-ratio and significance codes	Mouse	Dwell-2x	Dwell-4x	Push-2x	Push-4x
Mouse	-	148.6***	93.2***	142.3***	99.3***
Dwell-2x	148.6***	-	6.4	0.1	5.0
Dwell-4x	93.2***	6.4	-	5.2	0.1
Push-2x	142.3***	0.1	5.2	-	3.9
Push-4x	99.3***	5.0	0.1	3.9	-

These results show that the mouse is significantly more accurate (4.93 pixels) than all the other gesture-based techniques. There is no significant difference between the gesture-based conditions, however figure 18 indicates a slight advantage for both '4x' conditions (Dwell-4x = 11.37 pixels, Push-4x = 12.05 pixels) as opposed to both '2x' conditions (Dwell-2x = 13.62 pixels, Push-2x = 13.41 pixels). A repeated measures analysis of variance on distance, with trial and condition as main effects and subject as blocking factor revealed no significant effect ($F_1 = 2.52$, $p > 0.05$), indicating that there was no learning effect for accuracy over these three trials.

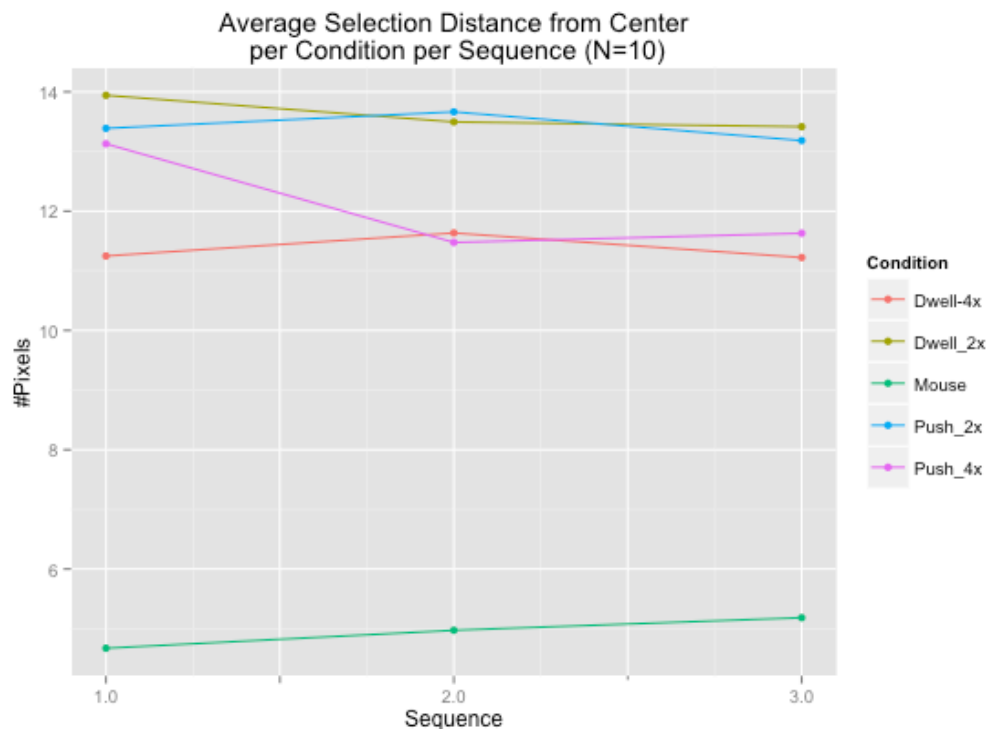


Figure 19. Average distance from the centre of the target per condition over three different sequences. The y-axis indicates the average number of pixels from the centre per selection.

Number of errors

Figure 20. shows the average number of errors per condition per sequence. A repeated measures analysis of variance, with condition as main effect and subject as blocking factor, indicated a significant effect between conditions ($F_4 = 3.82$, $p = 0.01$). To find out which condition(s) differed significantly, linear contrasts were formed, after which an F ratio was obtained, corrected for Type-I errors and finally a Tukey probability was calculated. The following results were obtained (Table 9.):

Table 9. Results of the post-hoc tests on number of errors obtained by forming linear contrasts, computing an F ratio, correcting for Type-I errors and finally computing a Tukey probability. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

F-ratio and significance codes	Mouse	Dwell-2x	Dwell-4x	Push-2x	Push-4x
Mouse	-	16.4**	0.4	15.9**	1.2
Dwell-2x	16.4**	-	11.6*	0.0	8.7*
Dwell-4x	0.4	11.6*	-	11.3*	0.2
Push-2x	15.9**	0.0	11.3*	-	8.3*
Push-4x	1.2	8.7*	0.2	8.3*	-

These results show that the mouse causes significantly less miss-clicks compared to both '2x' gesture conditions, while this is not the case for the '4x' gesture conditions. Furthermore Dwell-2x and Push-2x are both significantly more error-prone compared to both '4x' conditions. There is no significant difference however between both '4x' and between both '2x' conditions. A repeated measures analysis of variance on error, with trial and condition as main effects and subject as blocking factor revealed no significant effect ($F_1 = 0.97$, $p = 0.002$), indicating that there was no learning effect for number of errors over these three trials.

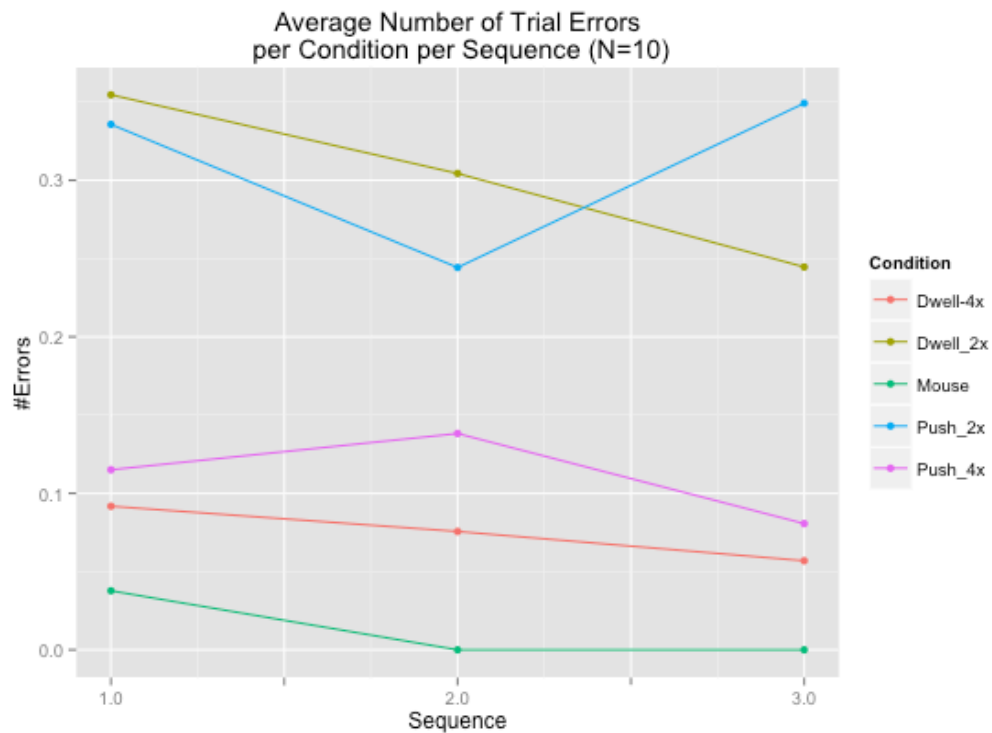


Figure 20. Average number of selections outside the target (errors) per condition over three different sequences. The y-axis indicates the average proportion of errors per selection

Throughput

Figure 21. shows the average throughput per condition per sequence in bits per second. A repeated measures analysis of variance, with condition as main effect and subject as blocking factor, indicated a significant effect between conditions ($F_4 = 69.11$, $p < 0.001$). To find out which condition(s) differed significantly, linear contrasts were formed, after which an F ratio was obtained, corrected for Type-I errors and finally a Tukey probability was calculated. The following results were obtained (Table 10):

Table 10. Results of the post-hoc tests on throughput obtained by forming linear contrasts, computing an F ratio, correcting for Type-I errors and finally computing a Tukey probability. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

F-ratio and significance codes	Mouse	Dwell-2x	Dwell-4x	Push-2x	Push-4x
Mouse	-	980.9***	845.0***	942.9***	851.3***
Dwell-2x	980.9***	-	5.1	0.4	4.6
Dwell-4x	845.0***	5.1	-	2.7	0.0
Push-2x	942.9***	0.4	2.7	-	2.3
Push-4x	851.3***	4.6	0.0	2.3	-

These results show that the mouse has a significantly higher throughput (average of 5.97 bits/s) than all the other gesture-based techniques, of which Dwell-4x (1.52 bits/s) and Push (1.51 bits/s) scored the highest, however there is no significant difference between the gesture-based conditions. A repeated measures analysis of variance on throughput, with trial and condition as main effects and subject as blocking factor revealed a significant effect between trials ($F_1 = 6.73$, $p = 0.01$), indicating that there was a learning effect for average throughput over these three trials, in which participants achieved a higher throughput over time (1st trial: 1.74 bits/s, 2nd trial: 1.81 bits/s, 3rd trial: 1.81 bits/s on average over all conditions).

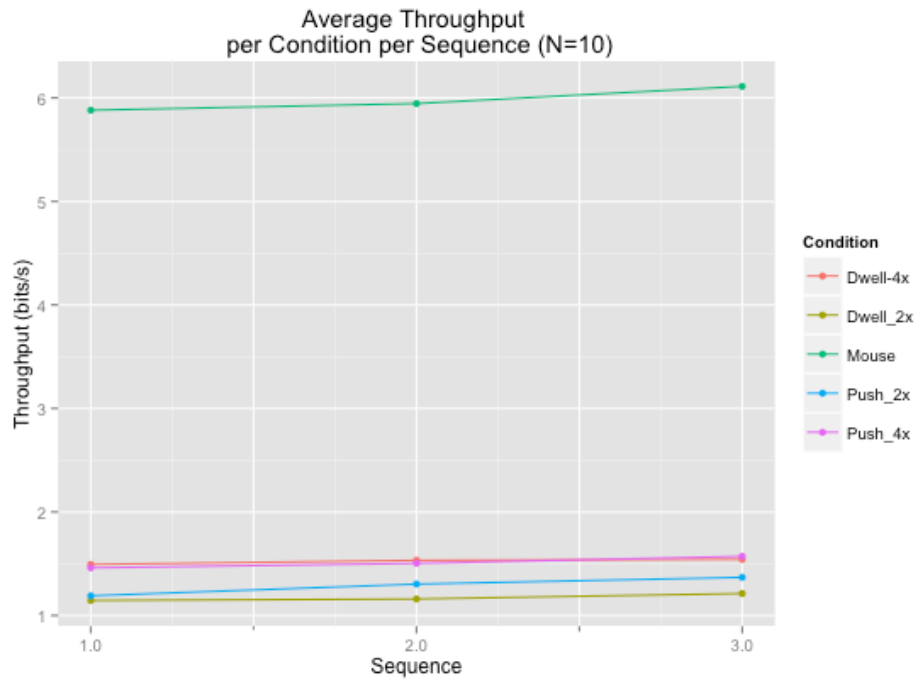


Figure 21. Average throughput per condition over three different sequences. The y-axis indicates throughput in bits per second.

6.3 Usability results

Gesture-based interaction questionnaire response

Participants were also asked to fill out a questionnaire concerning selection technique specific questions. This questionnaire deals with eight alternating negative and positive formulated statements. The results have been corrected for positive outcome, and can be found in Figure 22. Participants are most positive in general about the Mouse. Of the gesture-based interaction techniques, Dwell-4x is considered the most usable on nearly all properties, followed by Push-4X. Subsequently Dwell-2x and Push-2x score lowest on usability.

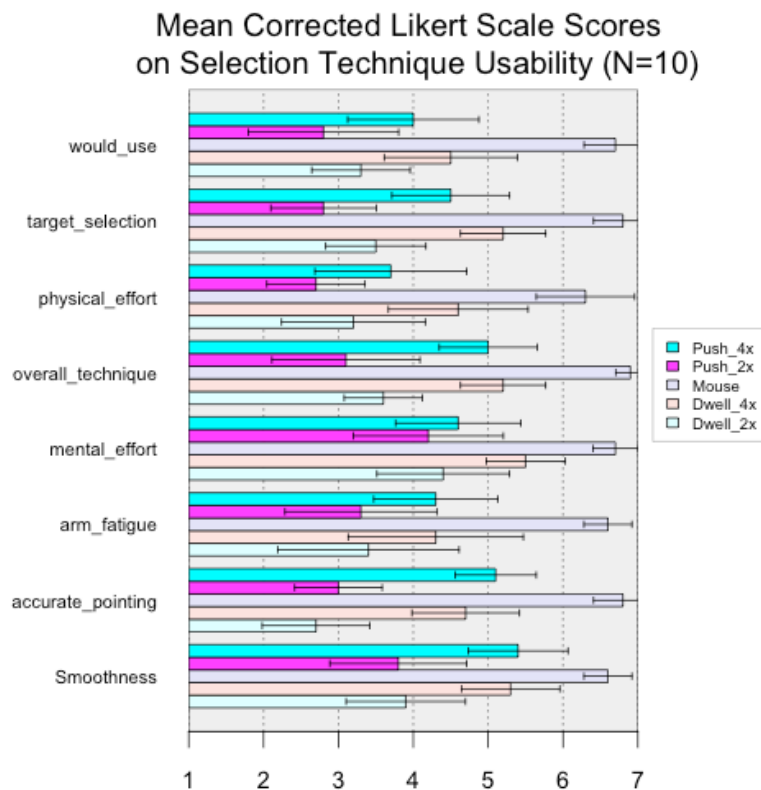


Figure 22. Stacked bar chart depicting the Likert scale responses concerning the usability on five selection techniques. The centre line depicts the mean value and error bars indicate a 95% confidence interval of the standard error.

6.4 Summary

The most important results of the selection technique evaluation can be summarized as follows: the mouse was considered significantly superior to the gesture-based interaction techniques on all performance measures (task-time, accuracy, number of errors and throughput). Of the gesture-based techniques, Dwell-4x and Push-4x performed significantly better on task time and number of errors compared to Dwell-2x and Push-4x, although no significant effect was found for accuracy and throughput. Furthermore no effect was found which indicated which of the two gesture techniques (Dwell and Push in general) performed better. Although there are small (and no significant) differences in throughput between the different gesture-based techniques, there are much larger differences in usability ratings however. From the Usability data it appears that besides the Mouse, Dwell-4x is preferred as selection technique followed by Push-4x. These two techniques received average positive results while Dwell-2x and Push-2x are much less preferred and even score negatively on certain important aspects such as physical efforts and perceived accuracy.

Chapter 7.

Usability recommendations

This chapter discusses usability issues of the test-case system based on the results of the usability evaluation in chapter 5 and chapter 6, and suggests possible usability improvements. Furthermore general guidelines for future developments in gesture-based interaction systems for the operating room are proposed. It is important to note that suggested solutions for the addressed usability issues are either based on the results of the experiments or relevant literature or in the case both are absent on the view of the author.

7.1 Observed test case usability issues

The exploratory research in chapter 4 and the user-based usability evaluation of the MITO PACS viewer in chapter 5 yielded interesting points of possible usability improvements. During the usability evaluation participants were given the possibility to give feedback on the test case system and indicate whether they missed certain functionality and other issues they would like to see improved (for detailed results see 5.3). These improvements are subdivided into the following categories: functional improvements, gesture improvements, interface improvements and miscellaneous improvements.

Functional improvements

Functional improvements for the MITO software include the option to switch between patient studies. This is very important because surgeons often want to look at different studies or old and new image studies of the patient. Also it should be possible to switch between image source type, switching between CT, MR and X-ray scans for example. Often surgeons would also like to have multiple viewports on the same screen, so that they can compare different studies or different viewing planes of the same study (in the case of 3D image reconstruction). Ideally MITO should also be able to display 3D image reconstructions and allow for 3D interaction, but this is not a must and would require a more complex interaction style in three dimensions.

Furthermore surgeons would like an option to activate/deactivate interaction with the system. Currently MITO starts recognizing gestures when the user's arms are outstretched to a certain extent, but surgeons would rather see a more definitive way of pausing interaction so that the system does not accidentally recognizes movements as meaningful gestures and gives the surgeon a higher sense of being in control. From the earlier exploratory research in chapter 4 respondents suggested the use of voice-activation, an onscreen pause button or a distinct pause gesture/posture. Of these possibilities voice-activation is least desirable in the context of the current implementation and other usability improvements. Voice-control is difficult to implement due to the noisy operating room environment, and confusion might arise in this highly cooperative working environment when spoken commands are expressed. A pause button is more in line with the existing interface and does not have any known negative influences on the interaction experience.

Surgeons also suggested that the system should have the option to undo a certain action. MITO has the possibility to clear a selection while measuring, but this does not currently act as an 'undo' gesture for all other functionality. Furthermore an option to reset the image to the default values would be appreciated.

Less frequently mentioned but nonetheless possible valuable suggestions include the addition of optional patient information such as blood values onscreen, motion picture support (especially useful for angiographic studies), integration of intra-operative navigation and operating room light control. This last suggestion would integrate a desirable extra feature, allowing surgeons to directly control the lighting so that again they do not need to ask anyone to do this for them. Interestingly Hartmann and Schlaefer

are looking into ‘the feasibility of touch-less control of operating room lights’ (2013), and concluded that it is indeed feasible and that their tested system is “intuitive even for inexperienced users”.

Gesture improvements

In general users would like to see a more continuous transition from one function to another. At present users have to deactivate the interaction with the system before a different function can be selected, by moving their hand in and out of the activation area. It would be more intuitive to be able to switch from zooming in on the image to rotating the image without interruption for example. This type of continuous interaction is common on smartphones and tablet computers. Integrating continuous interaction would be more in line with the experience of the user and thus their mental model. Furthermore surgeons indicate that gestures should preferably be made below the chest area, because this is considered the most sterile area of their gown and interaction space (Bible, Biswas, Whang, Simpson, & Grauer, 2009).

Selection in a clinical image viewer plays an important role. Considering that this is a major challenge in gesture-based interaction, due to the absence of physical buttons and feedback, this should receive considerable usability attention. In the current MITO gesture-based interface, selection only takes place when conducting measurements or selecting a region of interest, and proved to be less accurate than asking an assistant and participants indicated their discontent with the usability of the used technique (Chapter 5). The selection techniques as tested in Chapter 6 offer possible solutions, the ‘Dwell’ technique is considered the most usable of the tested gesture-based selection techniques, but Dwell has the undesirable property of accidentally selecting unwanted targets, which negatively influences the degree of the users’ control. The tested ‘Push’ technique has the undesirable property that when a user selects by pushing his non-dominant hand forward his or her pointing hand, which controls the cursor, moves slightly which is undesirable. One solution is to combine both techniques, by activating the Dwell technique when the non-dominant hand is used as trigger. This could for example be accomplished by holding the non-dominant hand upwards when wanting to conduct a selection. This would give the user the desirable property of control of the Push selection technique while maintaining the accuracy of the Dwell selection technique.

As mentioned in the aforementioned functional improvements, participants indicated the need for a way to pause the system. This could be implemented by assuming a certain posture for a short amount of time or by selecting a pause button onscreen. Switching between different studies or image types could be accomplished by swiping in the direction of the wanted screen. Similar to how navigation is generally implemented in touch-screen devices. A possible suitable gesture for bringing up patient information could be to act out a pull-down gesture (as if pulling a physical label). For altering the light conditions in the operating room, a single-handed turning gesture might be suitable (as if turning a physical knob).

Some modifications of the exiting gestures should also greatly enhance the usability. A useful modification according to one participant would be to adjust the animation gesture so that it allows for gesturing in the viewing plane of the image study (e.g. transversal, sagittal or coronal) instead of gesturing in one predefined horizontal plane, as is the case in the tested system. Other participants indicated that the current gesture for changing the window-width and window-level should be more similar, now they require contradictory movements, which lead to confusion as can be seen in the number of wrong gestures issued in Chapter 6. These results also indicated that the ‘undo selection’ gesture of waving the dominant hand and animating through the patient study, should work for both hands because participants tended to forget which hand was needed to execute this function during the usability evaluation.

Surgeons would like to be able to use the current two-handed gestures as well as small one-handed gestures, this is currently not possible due to the limited resolution of the

Kinect but future developments in successors are promising¹. This would allow for smaller and finer movements, which are essential in maintaining a sterile environment as the hands do not have to move outside of the most sterile working environment of the surgeon, furthermore the risk of accidentally touching a colleague or any other physical object in the vicinity of the surgeon is minimized. Another advantage of one-handed gestures is that surgeons would be able to hold an instrument in one hand while interacting with the other. Finally one participant suggested an improvement, needed due to a sterility issue, to the current gesture for extracting a region of interest. Currently a user has to cross his or her arms, which is not advisable during surgery due to possible crossover of pathogens and thus higher risk of endangering sterility. A possible better gesture would be to implement a cutting-like gesture by moving both open hands forward in a slicing motion for example.

Interface improvements

The MITO interface can be improved by extending the current interface with a screen to switch between different studies. This could possibly be visualized by using large thumbnails of preselected (pre-operative) studies. Also an extra screen could be added in which surgeons can access miscellaneous settings such as operating room lighting or important patient information.

The current main gesture-based interface should have an added pause button so that surgeons can indicate whether they want the system to be tracking their movements or not. Also there should be a possibility for multiple viewports, which can either contain different studies or different views of the same study.

When gesture-based selection is optimized as discussed earlier, a useful addition would be icons that visually represent the functionality needed to interact with the interface in the standard WIMP-style. This is useful for novice users, who are not yet acquainted with the gestures (which have to be learned) but also for users who have forgotten a gesture belonging to a certain function they want to use. This allows users to always be in control; novice users can use the more time costly but always present interface widgets to reach their intended goals while expert users can achieve these goals more efficient due to unique gestures, similar to using short keys in nearly all modern software.

System feedback is especially important in deviceless interaction. System feedback could be improved by adding a small video stream of the view of the Kinect and descriptive labels indicating the current state of tracking. Also the main screen could possibly indicate more clearly which state the system is in, whether it is activated or deactivated, which user is being tracked and which gesture is currently being recognized. Also the displayed DICOM information of the currently viewed study could be displayed in a designated undisturbed area onscreen. The current location onscreen is transparent and thus its readability is influenced by the background patient image. This is also the case for measurement information, which is now depicted in the study image. This should be made more readable or get a separate area onscreen outside the viewport. Finally angular information should also be added when selecting line segments and regions of interest.

Miscellaneous improvements

The calibration step was considered very tedious and frustrating. This could be merely a technical issue, but priority should be given to a fast and simple calibration process. Also the software should save the calibration information, because currently when the main gesture-based interface is closed (to select a different patient study for example) all information is lost and the whole calibration process has to be restarted, which is very inefficient.

¹ <http://mashable.com/2013/05/22/xbox-one-kinect-heartbeat/>

7.2 Guidelines for gesture-based interaction systems for operating room purposes

In order to give a more general account of gesture-based interaction in the operating room, the following guidelines are proposed for future systems. These are partially based on the findings of the test case usability improvements suggested above and partially on relevant literature.

Technical design considerations

The following suggestions for the technical aspects of a gesture-based medical image viewer should be considered.

Selection

Ideally selection would take place by using the Thumbtrigger gesture (Grossman et al., 2004). This is not yet possible with the current affordable gesture-recognition devices such as the Kinect due to their low spatial resolution, especially when the user is situated at a relatively large distance from the camera (1-3 meters). For now the best compromise is to use the 'Push' interaction technique to select GUI elements, while a combination of 'Push' and 'Dwell' (see chapter 6) is advised for conducting precise measurements on the medical image. Users could possibly hold up their non-dominant hand when wanting to conduct a measurement while the dominant hand has to keep still within the image for about 1.5 seconds to trigger a selection. Furthermore it is interesting to note that the importance of pointing precision depends on the 'granularity' of information being referenced (Mentis, O'Hara, Sellen, & Trivedi, 2012). The proposed combination of point and selection techniques would take this notion into account because it offers the possibility to create different modes of interaction that allow the user to switch between fine- or course-grained pointing on screen. Namely users could have course-grained control while navigating menu items and more fine-grained control while pointing within the image viewport.

Functionality

For the functionality of a gesture-controlled PACS viewer, the taxonomy in Figure 23. is proposed. This taxonomy consists of a 'General gesture interaction' part and a 'Specific gesture interaction' part. This division is chosen because it allows novice users to be able to interact with the system by using the simple general interaction gestures, but it also allows more experienced users to use more specific "shortcut" gestures. More importantly this division provides a fallback option whenever the user forgets a certain gesture during interaction. Such a fallback option is recommended in critical situation in which the reliability of such a system is of utmost importance (Artinger et al., 2011). Furthermore the previously suggested distinction in selection techniques between fine-grained and course-grained cursor control is also depicted in the same figure.

The 'General gesture interaction' refers to a mode in which the user can control the cursor and select GUI elements in a WIMP style. Selection takes place by pointing at the menu item of interest and pushing with the non-dominant hand. This way of selecting GUI elements is preferred over the Dwell selection technique because it does not accidentally select unwanted menu items.

The 'Specific gesture interaction' refers to a mode in which gesturing is used as main interaction method, as was the case in the testes MITO system. In this mode which each function is associated with a specific and intuitive gesture. These gestures can be regarded as shortcut 'keys'. Notably in order to conduct a selection in a study image, a combination of the Push and Dwell technique is advised. The user holds his non-dominant hand up and activates the measure state of the system, after which he or she can select two or more points of interest by holding the dominant hand steady for about 1.5 seconds for each selection. Although no significant difference was found in accuracy between Dwell and Push in Chapter 6, Dwell showed slightly higher accuracy scores and lower error scores and was considered most usable by participants.

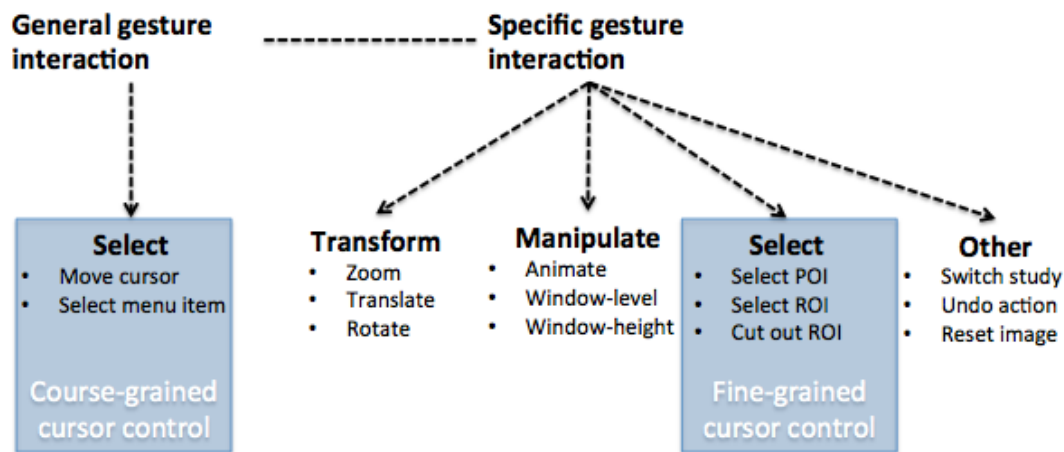


Figure 23. An overview of needed functionality for a gesture-based interaction system to control a medical image viewing application.

Activation/deactivation

In order to activate or deactivate the gesture interaction with the interface, the general gesture interaction mode could be used to control the cursor, in which a user can select a 'pause' button onscreen. This will deactivate the specific gesture interaction mode and freeze the onscreen functions but still allows the user to move the cursor so that the 'pause' button can be selected again and interaction can be resumed.

Calibration

If a calibration step is needed (depending on the used gesture detection technique), it should be performed beforehand and the system should be able to remember these settings during the whole session. Also users should be able to load and save calibration settings.

Training

New users can learn the systems' basic functionality with minor instructions in its suggested form. The 'General gesture interaction' mode only requires the user to know how to move the cursor and how to perform a selection gesture. The 'Specific gesture interaction' mode however requires a training phase in which the user will have to be instructed which gestures belong to which functions. This could be achieved by displaying a video instruction accompanied by text, which is the best way of instructing people for training gestural interactive systems according to Fothergill et al. (2012). A more interactive approach would be to combine these videos with an interactive tutorial in which the user has to act-out the displayed gestures in order to complete prototypical tasks. During these trials the user can be provided with interactive feedback.

Gestures

Table 11. describes possible useful and feasible gestures for the functions defined in figure 22. This table is divided in a column describing currently feasible gestures, which can be detected by current affordable gesture-detection cameras such as the Kinect, and a column describing possible one-handed alternatives when future developments offer higher resolution gesture detection.

Table 11. Overview of suggested gestures. A division is made between current feasible gestures and possible future gestures when gesture detection techniques are improved.

Function	Currently feasible gesture	Possible future gesture
Move cursor	Pointing with dominant hand using ray-casting technique	-
Select menu item	Pushing forward with non-dominant hand using Push technique described in Chapter 6	ThumbTrigger gesture (Grossman et al., 2004)
Animate	Moving non-dominant opened hand palm in the view direction of the images	-
Zoom	Moving both opened palms outward to zoom in and inward to zoom out	Pinch and zoom gesture using two fingers instead of two hands
Translate	Moving both opened palms in the direction of the wanted location	-
Rotate	Rotating both closed hands as if turning a steering wheel	Rotating two fingers instead of two hands
Window-level	Moving one opened hand and one closed hand simultaneously in the vertical plane	-
Window-height	Moving one opened hand and one closed hand simultaneously in the horizontal plane	-
Select point(s) of interest	Holding non-dominant hand opened while pointing with the dominant hand inside the image viewport and triggering a selection by holding the pointing hand still for a predefined amount of time	ThumbTrigger gesture (Grossman et al., 2004)
Cut-out region of interest	Bringing both opened hands forward in a slicing motion	Circular gesture with finger of dominant hand
Switch between screens	Swiping dominant hand in off-direction of target screen	-
Undo action	Making a waving motion with one of both hands	-
Reset image	Making a waving motion with both hands	-

Placement

From the questionnaire in Chapter 4 and from the test case usability evaluation in Chapter 5 it is clear that surgeons have a preference for having a gesture-based interaction system implemented beneath a separate monitor on a moveable arm so that it can be placed in the line of sight combined with the same image displayed on a large monitor on the wall so that everyone in the operating room can see the relevant patient images.

Extra functionality

Surgeons have indicated that they would like to be able to control operating room lighting. Such a feature should be incorporated in a separate screen in which the surgeon can select the relevant light and alter the lighting strength and direction. Such added functionality could greatly enhance the experience and persuade more users to use such a system.

Interface design considerations

The following guidelines concern the graphical user interface of the gesture-based medical image viewer, which is informed by the required functionality described above.

Interface layout

For the interface of such a system it is important to take into account the importance of image selection, medical image exploration and other interfaces (such as OR lighting control). All these options cannot be integrated clearly in one single screen, so it is advised to implement multiple screens between which the user can switch. Furthermore it is important to take desktop interface idioms into account when using gestural interfaces (Wobbrock, Morris, & Wilson, 2009) especially with the proposed division in general and specific gesture interaction mode in mind. A proposal for a gesture-controlled PACS viewer can be found in Appendix D.1.

Feedback

Due to the lack of kinaesthetic feedback it is important to focus on other means of feedback. First of all a feedback screen should be implemented that displays the current state of the tracking system. This makes the user aware whether he or she is being tracked and which gestures are being detected. Feedback can be provided textually, but also different colours can be used to signal the state to the user. Furthermore it is important to provide users with feedback concerning the gestures they are using. For example so-called 'ghost' hands could be displayed 'behind' the screen indicating the proximity of the hands to a 'virtual' touchscreen, the closer the users' hands are (visually) to the screen the larger the effect of interaction is. Finally the interface could add visual elements to the cursor or the 'ghost' hands (Artinger et al., 2011) indicating which function is activated and a graphical indication of the parameter limits (such as window-level value) or the direction of the gesture (such as when rotating).

Chapter 8.

Discussion

The main research question was to find out whether gesture-based interaction can serve as a more natural and usable way of viewing and manipulating medical images as opposed to asking an assistant to control the traditional keyboard and mouse, which is time consuming and hardly ever produces the desired result.

In an ideal situation, a surgeon would be able to access a patient's imaging study and interact with these images by issuing intuitive gestures towards the screen. He or she would never have to leave the sterile surgical work field and would never have to explain to an assistant what he or she would like to see. Gesture-based interaction would thus offer direct and sterile access to important functions needed to create a clearer view of the affected area, normally accessed by mouse and keyboard. Furthermore a surgeon would be able to gesture with small one-handed gestures, so that movement is confined to a safe and sterile interaction area while the other hand can continue with important surgical procedures.

Due to the lack of usability research conducted on gesture-based interaction techniques in the operating room, a thorough usability study was conducted to find out whether surgeons regard this type of interaction desirable, if it is suited for real operating room usage and how it can possibly be improved.

This chapter discusses the main findings of this study. First the results of this thesis are discussed in light of the research objectives, after which the implications of this thesis for the work field of the surgeon, the limitations of this study and suggestions for future research are discussed.

8.1 Summary of results

The main research question of this study was to find out how a gesture-based interaction technique using the Kinect can be implemented for operating room purposes. This broad research question was broken down into three sub-questions, which will be discussed separately: What do surgeons expect from gesture-based interaction? Can the Kinect serve as a better way of interacting with medical images as opposed to asking an assistant? How can gesture-based interaction in the OR be improved?

What do surgeons expect from gesture-based interaction?

This question was answered by the results obtained from the exploratory study described in Chapter 4. First of all the conducted interviews and operating room observations clearly indicated the desirability of a gesture-controlled medical image viewer. An interesting observation was that surgeons selected the most important image before surgery begins, and often did not interact with these images during surgery. This is striking because in the questionnaire distributed under surgeons, the majority of respondents indicated that they would want to interact with the patient images more often than is currently the case. Also they regarded a gesture-controlled viewer as a promising alternative to the current operating room situation.

Secondly the questionnaire gave insights in the most frequently used functions of the clinical image viewer, which are mainly basic functions such as animating through a patient study, zooming in and out, conducting measurements and changing contrast values. This information is important for specifying the functions that a gesture-based system should at least have incorporated, but also which functionality is possibly not necessary; too much functionality will adversely affect the usability of a gesture-based interface, due to the limited amount of gestures that can be remembered. The questionnaire also gave important insights for the setup of a test case usability evaluation, such as the preference for using a moveable arm with attached monitor in combination with the gesture-based interface but also that using a footswitch as activation/deactivation method of such a system was less preferred.

Can the Kinect serve as a better way of interacting with medical images as opposed to asking an assistant?

This question was answered by means of the user based usability evaluation of the gesture controlled medical image viewer “MITO”, which was conducted in Chapter 5. Usability was assessed by measuring several performance measures such as time and accuracy while participants performed prototypical tasks, such as locating a tumor and conducting measurements on a real patient study using the gesture-based test case system. Also a control condition was included in which the participant had to ask an assistant to conduct the same tasks. Furthermore participants’ subjective responses on the usability of the system were measured after the experiment.

Performance results indicated that the control condition, in which the surgeon directed an assistant to complete the tasks, is 1.25 times faster than using the tested gesture-controlled system. However a trained expert user was 1.72 times faster than the control condition, indicating that experience is clearly beneficial for the gesture-controlled system. In the previously mentioned study (see Chapter 2.3) by Ebert et al. (2011), participants required 1.4x more time to recreate images in the gesture condition as opposed to the control condition, in which the participant controlled the mouse/keyboard directly. Although the control condition is fundamentally different, the results in this study are in line with Ebert et al’s study: in this case participants in the gesture-condition required 1.25x more time to complete the tasks compared to asking an assistant, whereas participants in Ebert et al’s study required 1.4x more time compared to the easier control condition.

The accuracy of the test case was in some cases similar to the results obtained by asking an assistant in the control condition, except for changing window-level, window-width and conducting measurements, where the gesture condition was less accurate than the asking an assistant. Changing the window-level and width were rather difficult tasks, it is hard to accurately replicate the exact contrast settings of an image, so these might not have been the most representative tasks and no further conclusions will be drawn. A trained expert user showed similar accuracy results, or even better compared to the control condition, again indicating that training is possibly beneficial. Furthermore the amount of incorrectly issued gestures by the participants is highest for animating and clearing a selection, but this is mainly due to confusion in these one-handed gestures of the participants in which hand they chose to use (the correct gesture was often issued, but with the wrong hand). Finally the amount of incorrectly recognized gestures by the system is highest for selection and clearing a selection, which might be due to technical issues of the system and largely explain the higher completion times and diminished accuracy values of the measurement tasks mentioned above.

The questionnaire showed that participants were generally very positive about the usability of the tested gesture system. For example, participants indicated that they preferred the gesture-based system to asking an assistant, and suspected that they would already like to use the tested system in the operating room. Furthermore participants were free to give feedback on several aspects of the tested system. Positive aspects of such a system are its direct control, feeling of independence and increased sterility. One important negative aspect however was its diminished accuracy. Furthermore interesting suggestions for future improvements were proposed, such as a method to pause interaction with the system, suggestions for more intuitive gestures and added functionality, such as switching between patient studies, bringing up important patient information and being able to undo a certain action.

Due to the diminished accuracy of the tested gesture-based system compared to asking an assistant, an extra study was conducted in Chapter 6, which looked more closely into the accuracy and selection time properties of Kinect based selection techniques. This study showed that currently popular selection techniques are 2.31 times less accurate and 2.99 times more time costly than the mouse (a throughput at best of 1.52 bits/s versus 5.92 bits/s of the mouse). Interestingly both high-resolution (4x) conditions show faster

selection times compared to both low-resolution (2x) conditions, while participants have to interact with their hands in a larger tracking area. This might be because participants are less careful in the higher resolution condition and thus move their hands faster towards the target, while in the lower resolution condition they might be afraid to 'shoot over' the target, which is inefficient and thus leads them to move more carefully and thus more slowly.

The results show that accuracy and speed clearly depend on the resolution of the depth-camera used, which is hopeful for future developments since this resolution will most likely increase in the next few years.

In general these results indicate that the Kinect and thus gesture-based interaction techniques in general could potentially serve as a better way of interacting with medical images. Although the accuracy for novice users is lower compared to the current condition in which the surgeon directs an assistant, training can possibly improve the task-times and accuracy. Nonetheless surgeons rate the usability of the tested system high, and would already prefer to use this system than asking an assistant. This can most likely be attributed to the surgeon's feeling of independence and being in direct control, which probably outweighs the current diminished performance properties.

How can gesture-based interaction in the OR be improved?

The results of the usability evaluation and performance evaluation of selection techniques left much room for usability improvements, which were proposed in Chapter 7. These improvements concerned suggestions for the test case system and guidelines for future gesture-based systems for operating room purposes.

Gesture based-interaction can be improved by making a clear distinction between a novice and more experienced user. One major suggestion concerns a division in two different point and selection techniques and an interface tailored to this division. There should be a general gesture interaction mode, which allows the user to take control of the cursor and select functionality onscreen. Also there should be a more specific gesture interaction mode that allows the user to interact with the system in a more 'natural' way and in which different gestures belong to different functions. This distinction allows each user to pick up basic gesture-based interaction and after a while allows them to become experts by using 'shortcut-gestures'. Furthermore it offers backup optionality for users who might have forgotten a gesture, which is an essential feature in such a critical working environment. In order to increase the accuracy of measurements in gesture-based interaction, a distinction in selection techniques is proposed in which there is a general selection technique used in the general interaction mode and a more accurate selection technique, which is used in the specific interaction mode.

8.2 Implications

This study is the first extensive usability study on the possibilities of using gesture-based technology in the operating room in order to interact with medical images of the patient. The results showed that surgeons are very positive about the current state of art gesture-based technologies and would like such techniques to be refined and implemented in the operating room in the near future. This will clearly change the current workflow of the surgeon in a more direct, efficient and enjoyable way, possibly saving valuable time, money but more importantly lives.

The empirical methodology used in this study has shown to be successful in detecting the needs of surgeons, indicating the appropriate performance measures, testing a gesture-based interaction system and finally determining usability improvements. Especially the user-based usability evaluation has shown to be efficient in testing a completely new technique and eliciting useful feedback from the domain users. This methodology is thus very useful and recommended for future innovative interaction systems in the operating room, which need to be tested on their usability and for which no model-based usability evaluation exists and for which inspection-based usability evaluation does not suffice.

Furthermore this study is a welcome addition to the generally slim amount of usability research on gesture-based interaction in general. It provides a practical methodology that can be applied to new practical gesture-based interaction innovations, which will certainly become more and more popular in the years to come.

8.3 Limitations

One limitation of this study is that of the thirteen participants in the user-based usability evaluation, eight participants were actual surgeons whereas the other five were medicine students, however statistical analysis did not point out any significant differences between these two groups. More surgeons would have been more representative for the experiment and might have elicited more usability issues, however Nielsen¹ advocates that in most cases five to ten users is enough to elicit 80 to 95 % of all usability issues.

Another limitation concerns the control condition, in which the participants directed an assistant to complete the tasks. However no real assistants were used, in this case the observer acted as the assistant who exactly followed the instructions of the participants and showed no own initiative. This might not be representative of real surgical assistants, who have varying experience in exploring medical images and often know what the surgeon is interested in. In this study however the 'assistant' resembles a worst-case scenario against which the gesture-based condition can be compared. If the gesture-based condition were faster than the worst-case 'assistant' condition, this would indicate that it would always be beneficial to the surgeon, which was the main interest of this thesis.

An expert user was used in the usability evaluation in order to see what the effect of experience is on the predefined performance measures. In this case only one expert user conducted the tasks in the usability evaluation three times. To be able to infer more about possible learning effects, it would be interesting to repeat this study over time on all participants to see whether they improve.

The tasks that were conducted in the usability evaluation were designed in such a way that they included all functionality that was implemented in the test case system, which complied to the essential functionality according to the surgeons in the exploratory research phase. However these tasks were defined rather strictly, in such a way that there was only one way to solve each subsequent subtask. This was useful for determining performance measures because participants were forced to complete the subtask in one particular way allowing for better comparison between the experimental and control condition, but did not allow the surgeon to interact freely and solve a higher-level task in his or her own way. Furthermore several subtasks were fairly difficult for the participant to recreate, such as changing the window-level and window-width, because it is hard to exactly match a target contrast setting. This finding was also reflected in the low accuracy results of these subtasks. For future studies it is advisable to take this into account, it would be interesting to repeat this study with several different patient studies and several different high-level tasks to get an even more realistic view of the usability of a gesture-based system compared to the current situation.

Finally, although the usability results are generally positive in favour of gesture-based interaction in the operating room, results need to be interpreted with caution, as gesture-based interaction could also be benefitting from a novelty effect, i.e., an increased interest in new technology. Further experiments with different systems and possibly during real surgical procedures, will have to demonstrate whether this is the case.

8.4 Future developments

As suggested above, it would be very interesting to test several gesture-based interaction systems in a comparative usability evaluation. This might yield interesting insights and possible overlooked usability issues.

¹ <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>

A lot of research is currently being conducted on intuitive gestures, especially on mobile touchscreen devices (Wobbrock et al., 2009) and to a lesser extent for device-less gesture interaction. It would be interesting to find out which gestures are considered most intuitive by surgeons for interacting with medical images. The method proposed by Nielsen et al. (2004) for developing intuitive and ergonomic gesture interfaces and the agreement analysis by Wobbrock et al. are interesting methods waiting to be applied to this specific domain (see Chapter 2.3).

Another interesting venture would be to look at the usability of so-called radial menus in gestural interfaces. This concept is starting to be used more often in modern touch screen interfaces (Artinger et al., 2011) and it would be interesting to see whether this could enhance the usability of device-less gesture based interfaces. Radial menus offer a way of bringing up a menu with context depended functionality, which can also be accompanied by instructions for example. Furthermore it provides the user with equally distributed menu items, which offers faster navigation as opposed to hierarchical menus, which might be far more useful for gesture-based menu selection.

One very interesting research direction would be to look into the possibilities of the recently announced Leap Motion¹ to act as a gesture-detection device. This system can recognize very small gestures of separate fingers nearby as opposed to the coarse gesture-detection of the tested Kinect. One challenge however would be to integrate the Leap Motion within the sterile working environment of the surgeon.

Finally the lack of model-based usability evaluation such as GOMS analysis or ACT-R models of gesture-based interaction calls for research in this area. Very recently a research group named Cogscen² has released a special variant of ACT-R, which contains an extension of the manual buffer that accounts for touch-screen interaction on the popular iPad. It would be interesting to modify this model to account for device-less gesture-based interaction, which might be challenging due to the complex nature of freeform gesture interaction.

¹ <https://www.leapmotion.com/>

² <http://cogscen.com/>

Chapter 9.

Conclusion

The results of this thesis indicate that the concept of gesture-based interaction using low-cost commercially available hardware, such as the Kinect, is feasible for operating room purposes. Although the accuracy is lower and execution times are slower compared to the current situation in which the surgeon directs an assistant, training and future technological innovations such as higher resolution depth-camera's can possibly improve the accuracy of gesture-based interaction. Nonetheless surgeons rate the usability of the gesture-based interaction high, and would already prefer to use the current state-of-the-art systems and be in direct control of the medical images, than keep on asking assistants to interact with them indirectly.

This study is the first study to knowledge to evaluate a gesture-based system in the operating room in such an extensive manner and to provide guidelines for future research. Other known studies hardly take the usability of their gesture-based technique into account. But by thoroughly studying the usability of such novel techniques in such specific domains, it can yield high benefits for surgeons and surgical assistant but also for companies producing these systems.

The methodology used in this study was successful in evaluating the usability of the test case gesture-controlled system and can therefore be used in future usability studies on other similar systems. Furthermore it was successful in eliciting suggestions for usability improvements, which were captured in guidelines for future designs.

The usability improvements and guidelines suggested in this thesis can be valuable for companies and healthcare institutions interested in gesture-based technology, because the suggestions allow them to make their systems more consistent with the requirements of the surgeon. This in turn benefits the surgeon by saving him or her time and taking away frustrations associated with asking an assistant to conduct the same task, which eventually benefits the health of the patient.

Future developments that take these usability recommendations into account in combination with improved hardware will eventually lead to a robust system that will most definitely be implemented in the operating room and offer the surgeon a positive and informative experience when wanting to inspect the patients' images.

Bibliography

- Artinger, E., Coskun, T., Schanzenbach, M., Echtler, F., Nestler, S., & Klinker, G. (2011). Exploring Multi-touch Gestures for Map Interaction in Mass Casualty Incidents. *Control*. Retrieved from <http://www.user.tu-berlin.de/komm/CD/paper/060722.pdf>
- Baudel, T., & Beaudouin-Lafon, M. (1992). An Interaction Model Designed For Hand Gesture Input. *LRI Res*, 772. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.21.5180>
- Bible, J. E., Biswas, D., Whang, P. G., Simpson, A. K., & Grauer, J. N. (2009). Which Regions of the Operating Gown Should be Considered Most Sterile? *Clinical Orthopaedics and Related Research*, 467(3), 825–830. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2635444&tool=pmcentrez&rendertype=abstract>
- Bigdelou, A., Schwarz, L., & Navab, N. (2012). An adaptive solution for intra-operative gesture-based human-machine interaction. *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces IUI 12* (p. 75). ACM Press. doi:10.1145/2166966.2166981
- Bolt, R. A. (1980). Put-that-there”: Voice and gesture at the graphics interface. *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, 14(3), 262–270. doi:10.1145/800250.807503
- Bongers, B. (2004). *Interaction with our electronic environment – an e-ecological approach to physical interface design*. Hogeschool van Utrecht. Retrieved from <http://bertbon.home.xs4all.nl/downloads/Cahier.pdf>
- Bowman, D. A., & Hodges, L. F. (1999). Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments. *Journal of Visual Languages Computing*, 10(1), 37–53. doi:10.1006/jvlc.1998.0111
- Brooke, J. (1996). SUS - A quick and dirty usability scale. (P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland, Eds.) *Usability evaluation in industry*, 189, 189–194. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:SUS+-+A+quick+and+dirty+usability+scale#0>
- Douglas, S. A., Kirkpatrick, A. E., & MacKenzie, I. S. (1999). Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. (M. W. Altom & M. G. Williams, Eds.) *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit CHI 99*, 15(May), 215–222. doi:10.1145/302979.303042
- Dumas, J. S. (2003). User-based evaluations. In J. Jacko & A. Sears (Eds.), *The humancomputer interaction handbook fundamentals evolving technologies and emerging applications* (Vol. User-based, pp. 1093–1117). L. Erlbaum Associates Inc. Retrieved from <http://portal.acm.org/citation.cfm?id=772141>
- Dumas, J. S., & Redish, J. C. (1999). *A practical guide to usability testing*. Star (Vol. Rev. ed., p. 404). Intellect. Retrieved from <http://www.amazon.com/Practical-Guide-Usability-Testing/dp/1841500208>
- Ebert, L. C., Hatch, G., Ampanozi, G., Thali, M. J., & Ross, S. (2011). You Can’t Touch This: Touch-free Navigation Through Radiological Images. *Surgical Innovation*. doi:10.1177/1553350611425508

- Fikkert, W. (2010). *Gesture Interaction at a Distance*. University of Twente.
- Fothergill, S., Mentis, H., Kohli, P., & Nowozin, S. (2012). Instructing people for training gestural interactive systems. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems CHI 12* (pp. 1737–1746). ACM Press. doi:10.1145/2207676.2208303
- Gallo, L., Placitelli, A. P., & Ciampi, M. Controller-free exploration of medical image data: Experiencing the Kinect. , 2011 24th International Symposium on ComputerBased Medical Systems CBMS 1–6 (2011). IEEE. doi:10.1109/CBMS.2011.5999138
- Grätzel, C., Fong, T., Grange, S., & Baur, C. (2004). A non-contact mouse for surgeon-computer interaction. *Technology and health care official journal of the European Society for Engineering and Medicine*, 12(3), 245–257. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15328453>
- Grossman, T., Wigdor, D., & Balakrishnan, R. (2004). Multi-finger gestural interaction with 3d volumetric displays. *Proceedings of the 17th annual ACM symposium on User interface software and technology*, 24(3), 61–70. doi:10.1145/1029632.1029644
- Hartmann, F., & Schlaefer, A. (2013). Feasibility of touch-less control of operating room lights. *International journal of computer assisted radiology and surgery*, 8(2), 259–68. doi:10.1007/s11548-012-0778-2
- Hinckley, K., Pausch, R., Proffitt, D., & Kassell, N. F. (1998). Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction*, 5(3), 260–302. doi:10.1145/292834.292849
- Holleis, P., Otto, F., Hussmann, H., & Schmidt, A. (2007). Keystroke-level model for advanced mobile phone interaction. *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 07*, 1505. doi:10.1145/1240624.1240851
- Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2), 79–102. doi:10.1016/j.ijhcs.2005.06.002
- Iso. (1998). Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. (I. S. Organization, Ed.) *International Organization for Standardization*. International Organization for Standardization. Retrieved from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16883
- Karam, M., & Schraefel, M. (2005). A taxonomy of Gestures in Human Computer Interaction. *ACM Transactions on ComputerHuman Interactions*, (ECSTR-IAM05-009), 1–45. Retrieved from <http://www.mendeley.com/research/taxonomy-gestures-human-computer-interaction-1/>
- Khoshelham, K., & Elberink, S. O. (2012). Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors (Peterboroug)*, 12(2), 1437–1454. doi:10.3390/s120201437
- Kurtenbach, G., & Hulteen, E. (1990). Gestures in Human-Computer Communication. In B. Laurel (Ed.), *The Art and Science of Interface Design* (pp. 309–317). Reading: Addison-Wesley Publishing Co. Retrieved from <http://www.autodeskresearch.com/publications/gestureshci>
- Lee, H.-K., & Kim, J.-H. (1998). Gesture spotting from continuous hand motion. *Pattern Recognition Letters*, 19(5-6), 513–520. doi:10.1016/S0167-8655(98)00027-0

- MacKenzie, I. S., Kauppinen, T., & Silfverberg, M. (2001). Accuracy measures for evaluating computer pointing devices. (Julie Jacko & A. Sears, Eds.) *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 01, Seattle, W*, 9–16. doi:10.1145/365024.365028
- Mentis, H. M., O'Hara, K., Sellen, A., & Trivedi, R. (2012). Interaction proxemics and image use in neurosurgery. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* (p. 927). New York, New York, USA: ACM Press. doi:10.1145/2207676.2208536
- Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., & Mackay, W. (2011). Mid-air pan-and-zoom on wall-sized displays. *Design* (pp. 177–186). ACM Press. doi:10.1145/1978942.1978969
- Nielsen, J. (1993). *Usability Engineering*. (J Nielsen, Ed.) *Usability Engineering* (Vol. 44, p. 362). Morgan Kaufmann. doi:10.1145/1508044.1508050
- Nielsen, Jakob, & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, & T. White (Eds.), *Transport* (Vol. 206, pp. 206–213). ACM. doi:10.1145/169059.169166
- Nielsen, M., Störring, M., Moeslund, T. B., & Granum, E. (2004). A procedure for developing intuitive and ergonomic gesture interfaces for HCI. (A. Gamurri & G. Volpe, Eds.) *GestureBased Communication in HumanComputer Interaction, 2915*, 409–420. doi:10.1007/978-3-540-24598-8_38
- Norman, D. A. (1991). Cognitive artifacts. (J. M. Carroll, Ed.) *Proceedings of the 6th ACM SIGCHI conference on Creativity cognition CC 07, 20(1991)*, 1–18. doi:10.1145/1254960.1255025
- Norman, D. A., Nielsen, J., & Group, N. N. (2010). Gestural Interfaces : A Step Backward In Usability. *interactions*, 17(5), 46–49. doi:10.1145/1836216.1836228
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3), 510–532. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8272468>
- Schultz, M., Gill, J., Zubairi, S., Huber, R., & Gordin, F. (2003). Bacterial contamination of computer keyboards in a teaching hospital. *Infection control and hospital epidemiology the official journal of the Society of Hospital Epidemiologists of America*, 24(4), 302–303. Retrieved from <http://www.jstor.org/stable/10.1086/502200>
- Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., et al. Real-time human pose recognition in parts from single depth images. , 2 CVPR 2011 1297–1304 (2011). IEEE. Retrieved from <http://ieeexplore.ieee.org/lpdocs/epico3/wrapper.htm?arnumber=5995316>
- T. Kipshagen, M. Graw, V. Tronnier, M. Bonsanto, U. G. Hofmann. (2009). *World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany*. (O. Dössel & W. C. Schlegel, Eds.) (Vol. 25/6, pp. 75 – 78). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-03906-5
- Van Ooijen, P. M. A. (2005). Physical and Logical network setup of a PACS with integrated AquariusNET server and Aquarius Workstation. *Clinical Case Studies in the Third-Dimension* (pp. 77–80). TeraRecon. Retrieved from http://www.terarecon.com/downloads/news/casestudy_vanooijen_aquariusandpacs.pdf

- Vatavu, R.-D. (2011). Point and click mediated interactions for large home entertainment displays. *Multimedia Tools and Applications*, 59(1), 1–16. doi:10.1007/s11042-010-0698-5
- Vogel, D., & Balakrishnan, R. (2004). Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In S. Feiner & J. A. Landay (Eds.), *Work* (Vol. 6, pp. 137–146). ACM. doi:10.1145/1029632.1029656
- Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. *Proceedings of the 18th annual ACM symposium on User interface software and technology UIST 05*, pp(3), 33. doi:10.1145/1095034.1095041
- Wachs, J. P., Stern, H. I., Edan, Y., Gillam, M., Handler, J., Feied, C., & Smith, M. (2008). A gesture-based tool for sterile browsing of radiology images. *Journal of the American Medical Informatics Association*, 15(3), 321–323. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2410001&tool=pmcentrez&rendertype=abstract>
- Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. (S. W. U. Jacob O Wobbrock University Of Washington, Ed.) *Proceedings of the 27th international conference on Human factors in computing systems CHI 09*, 1083. doi:10.1145/1518701.1518866
- Zhang, X., & Mackenzie, I. S. (2007). Evaluating Eye Tracking with ISO 9241 - Part 9. *Human-Computer Interaction*, 4552, 779–788. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=FD1001A57CE5D5EB2896A90E3BDCoA86?doi=10.1.1.72.8113&rep=rep1&type=pdf>
- Zimmerman, T. G., Smith, J. R., Paradiso, J. A., Allport, D., & Gershenfeld, N. (1995). Applying electric field sensing to human-computer interfaces. (D. V Keyson, M. Lou Maher, N. Streitz, A. Cheok, J. C. Augusto, R. Wichert, G. Englebienne, et al., Eds.) *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 95*, 35(May), 280–287. doi:10.1145/223904.223940

Appendices

Appendix A.1

Interview with Surgeons

- *Describe the current operating room situation with respect to the clinical image viewer and mouse and keyboard usage.*

The surgeons explained that they reinvestigate the radiological images before surgery starts to refresh their minds and to obtain a detailed internal representation of the case. Furthermore a surgical procedure plan is made and discussed with the rest of the team, this plan is illustrated with the radiological images.

- *Does the current situation cause delay or frustrations in the workflow? And is there a clear need for a system that can reduce or solve these problems?*

“Yes, the current situation often implies that we have to ask an assistant to navigate the images for us, and due to this indirect form of communication he or she never gets it exactly how we want”. Furthermore both surgeons clearly saw the benefits of a (gesture-controlled) system that can offer them a direct form of control over these images during surgery.

- *How do surgeons regard gesture-based interaction techniques for in the operating room?*

While demonstrating the test case system, they were very positive and clearly saw the benefits of such a system during surgery and they thought it would be interesting to test such a system.

- *Which functionality of the medical image viewer is used most often during surgery?*

According to both surgeons mainly basic functionality is used, such as zooming in and out as well as changing window-level and sometimes conducting measurements.

- *Is there a preference for two-dimensional or three-dimensional images?*

“Mainly two-dimensional images and sometimes three-dimensional”

- *What do surgeons think of voice-controlled systems?*

Both surgeons regarded voice control as a potentially good solution, but they are very skeptical about its success in the operating room due to the amount of (background) noise.

- *Which advantages and/or disadvantages does a gesture-based system bring to the operating room?*

Both surgeons regarded the direct control over the medical image viewer as major advantage as well as the enhancement of asepsis. Although they were concerned about large arm movements during surgery, which might in turn endanger asepsis.

Appendix A.2

Online questionnaire overview

Kinect-technologie in de operatiekamer?

Iedereen is inmiddels gewend aan de mogelijkheid om in de operatiekamer CT/MRI scans van patiënten te bekijken en te manipuleren door gebruik te maken van toetsenbord en muis.

Sinds kort zijn er innovaties in opkomst die andere, wellicht meer natuurlijke interactievormen mogelijk maken. Deze enquête richt zich op een alternatief dat gebruik maakt van handgebaren voor de manipulatie van medische beelden.

Met deze enquête willen wij graag inzicht krijgen in de houding van chirurgen ten aanzien van een dergelijk systeem. Ter illustratie wordt op de volgende pagina een filmpje getoond van het systeem dat op dit moment in het UMCG beschikbaar is.

Deze enquête bestaat uit 14 korte vragen en zal ongeveer 5 minuten in beslag nemen.

In deze enquête is uw anonimiteit gewaarborgd.

Start

www.thesistools.com

Kinect-technologie in de operatiekamer?

Het onderstaande filmpje toont de mogelijkheden van een handgestuurde interactiemethode die potentieel gebruikt zou kunnen worden in de operatiekamer:

De volgende vragen gaan over u en uw bekendheid met handgestuurde interactie

3.

Wat is uw specialisme binnen het UMCG?*

4.

Was u voor deze enquête bekend met het bedienen van een (spel)computer met behulp van handgebaren?

(Zoals de "Kinect" in het bovenstaande filmpje bijvoorbeeld)*

- ☐ Ja, ik heb dit zelf wel eens gebruikt
☐ Ik heb er wel eens over gehoord
☐ Nee, dit is nieuw voor mij

5.

Was u voor deze enquête al bekend met handgestuurde interactie, dus zonder toetsenbord en muis, in de operatiekamer?

(Zoals het systeem in het youtube filmpje bijvoorbeeld)*

- ☐ Ja, ik heb een dergelijk systeem wel eens gebruikt
☐ Ik heb er wel eens wat over gehoord
☐ Nee, dit is nieuw voor mij

De volgende vragen gaan over de huidige situatie in de operatiekamer

6.

Welke van de onderstaande functies (meerdere keuzes mogelijk) gebruikt u van de PACS viewer, om beelden van de patiënt te manipuleren gedurende of voor een operatie?

- ☐ In en uit zoomen
- ☐ Rotatie
- ☐ Het beeld verschuiven
- ☐ "Window-level" aanpassen
- ☐ Door een patient studie heen scrollen
- ☐ Met cursor gebieden aanwijzen
- ☐ Metingen doen
- ☐ Interesse gebied markeren
- ☐ Interesse gebied uitsnijden
- ☐ Andere studies van de patient bekijken
- ☐ Anders, namelijk:
- ☐ Anders, namelijk:
- ☐ Anders, namelijk:

7.

Directe interactie:

Als u tijdens een operatie uit de steriele zone stapt om medische beelden van de patiënt te manipuleren, hoe vaak doet u dit dan gemiddeld?

Maak een ruwe schatting:*

8.

Indirecte interactie:

Als u tijdens een operatie de beelden van de patiënt wilt manipuleren, hoe vaak vraagt u (gemiddeld genomen) een assistent om dit voor u te doen?

Maak een ruwe schatting:*

9.

Zou u in de ideale situatie, waarin het operatieproces nauwelijks onderbroken hoeft te worden, vaker direct de beelden van de patient willen kunnen manipuleren dan nu het geval is?*

- ☐ Ja
- ☐ Weet ik niet
- ☐ Nee

De volgende vragen gaan over uw eventuele voorkeuren met betrekking tot een handgestuurde interactiemethode

10.

Het voorbeeld systeem van het filmpje reageert op iedereen binnen het zichtveld van de camera.
Welke voorkeur zou u hebben voor het activeren/deactiveren van een dergelijk handgestuurd systeem?*

- ☐ Spraak-gestuurde commando's
- ☐ Voetpedaal
- ☐ Een assistent vragen
- ☐ Een bepaalde lichaamspositie aannemen
- ☐ Een (de)activatie knop aanwijzen
- ☐ Geen voorkeur

11.

Het voorbeeldsysteem werkt door middel van hand/lichaam herkenning en reageert op bepaalde handbewegingen

Uitzoomen is bijvoorbeeld gekoppeld aan: beide handen voor het lichaam uit elkaar bewegen.

Is het wenselijk dat een handgestuurd systeem, in bepaalde situaties, ook gebaren die met 1 hand gemaakt worden moet kunnen herkennen?*

- ☐ Ja
- ☐ Nee
- ☐ Weet ik niet

12.

Welke van onderstaande mogelijkheden heeft uw voorkeur als het gaat om het weergeven van de beelden van de patiënt.*

- ☐ Een groot TV scherm opgehangen aan de wand van de operatiekamer
- ☐ Een beweegbare arm met monitor die naar wens geplaatst kan worden
- ☐ Een combinatie van bovenstaande opties
- ☐ Anders, namelijk:
- ☐ Geen voorkeur

13.

Door wie zou een handgestuurd systeem bediend moeten/mogen worden?*

- ☐ Alleen de chirurg
- ☐ Chirurg & Assistenten
- ☐ Iedereen in de operatiekamer
- ☐ Anders, namelijk:
- ☐ Geen idee

14.

Ziet u een handgestuurde interactie-methode in de operatiekamer als meerwaarde ten opzichte van de huidige situatie?*

- ☐ Ja
☐ Misschien
☐ Nee

15.

Heeft u nog op-of-aanmerking met betrekking tot handgestuurde interactie of deze enquête?

16.

Mocht u interesse hebben in deelname aan dit onderzoek dan zouden wij graag met u in contact komen door hieronder uw emailadres in te vullen (dit emailadres zal alleen eenmalig voor contactlegging worden gebruikt).

Verstuur deze enquête

Wij danken u hartelijk voor uw participatie!

Vriendelijke groet,

Sebastiaan Stuij (Master student Human-Machine Communication)

Onder leiding van:

Dr. Ir. Peter van Ooijen (Medical Imaging Informatics, CMI-NEN)

Dr. Fokke Cnossen (Universitair Docent Cognitive Psychology & Ergonomics, Rijksuniversiteit Groningen)

In samenwerking met:

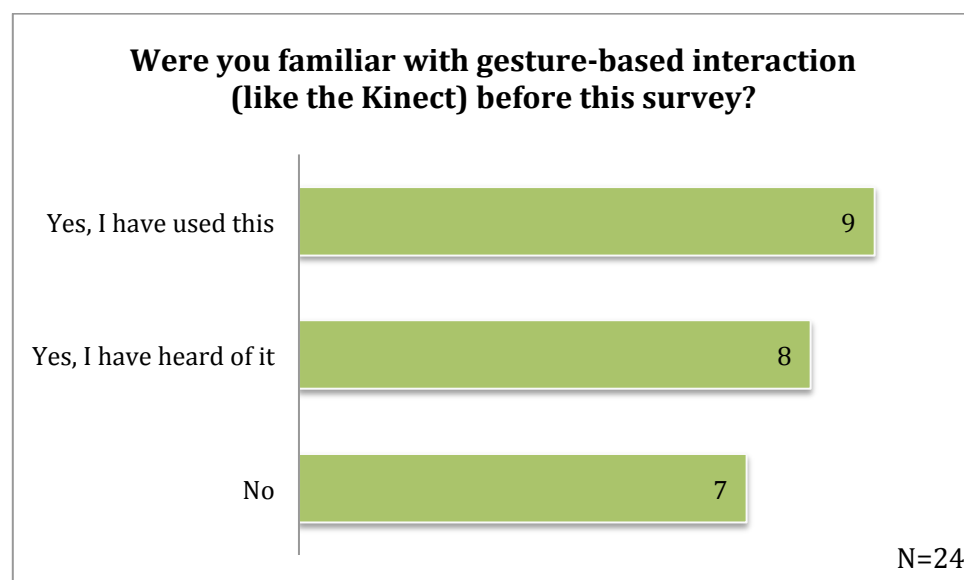
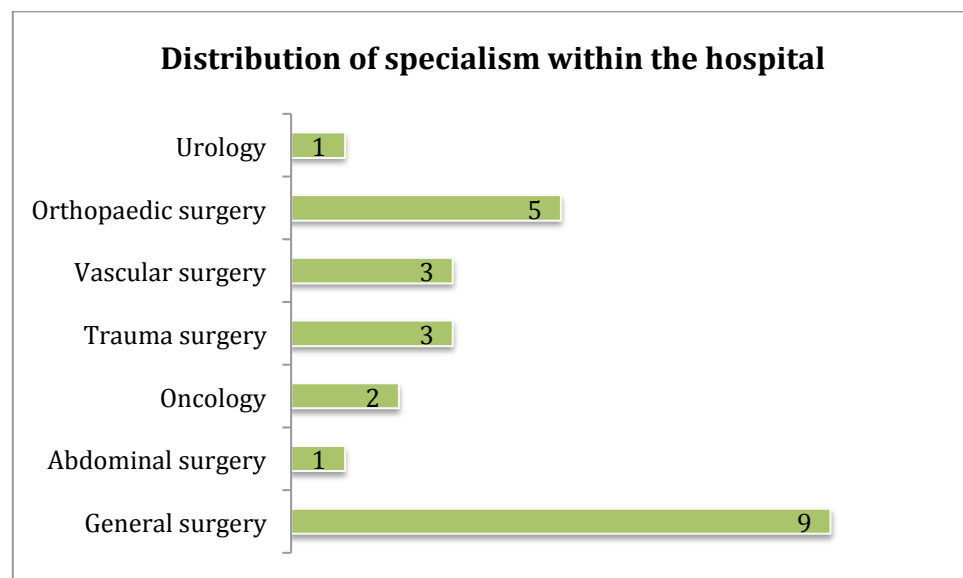
Drs. Henk ten Cate Hoedemaker & Jetse Goris (Wenckebach instituut UMCG)

Dr. Paul Jutte & Jasper Gerbers (Orthopedische chirurgie UMCG)

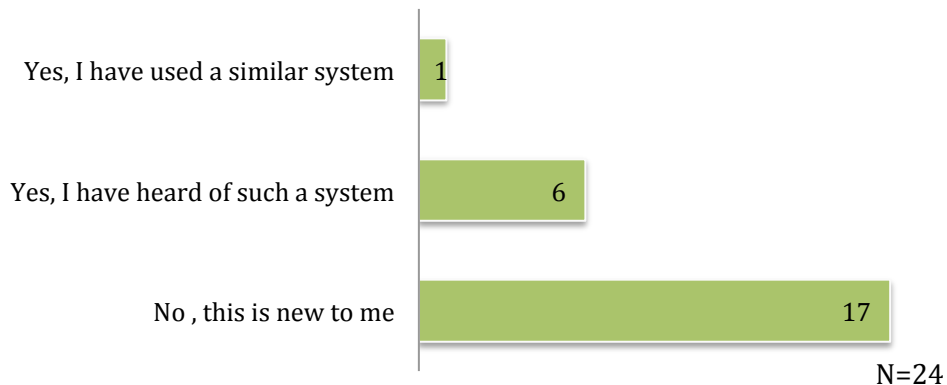
www.thesistools.com

Appendix A.3

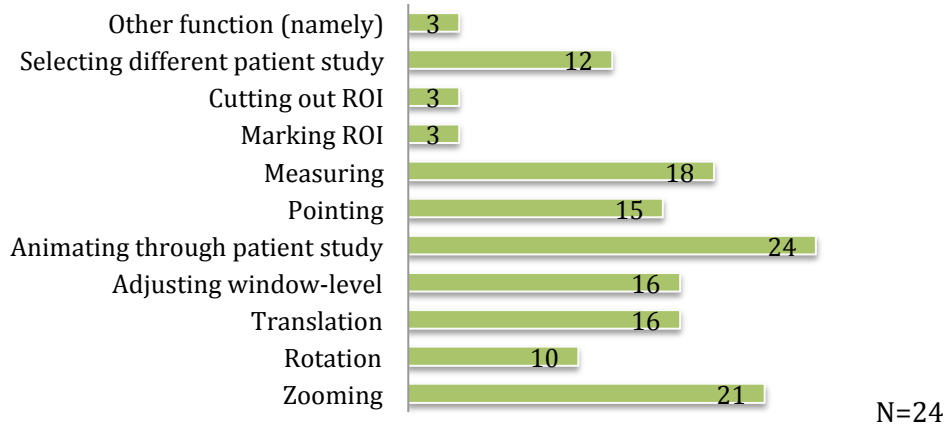
Online questionnaire results



Were you familiar with gesture-based interaction technology in the operating room before this survey?



Which functions of the PACS viewer do you use to view a patient study before or during surgery?



Other functionality:

- *Measuring angles*
- *Making 3d reconstructions*
- *Making central lumen line reconstructions*
- *Aligning multiple images in the same window*

Direct interaction:

When you step out of the sterile operating zone to interact with the patient study, how often do you do this on average per surgery?

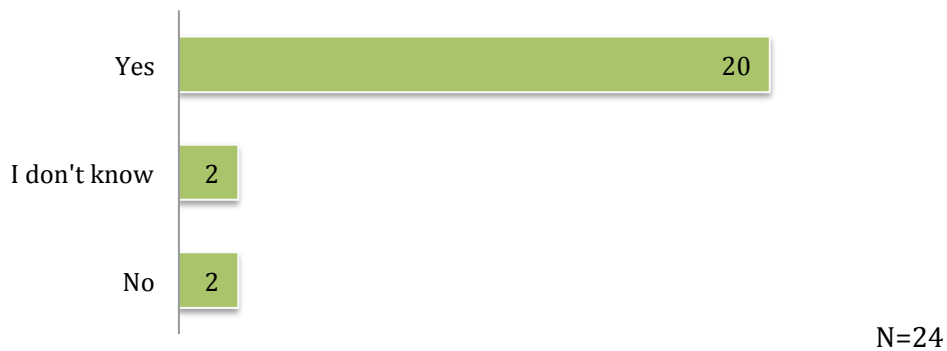
0.78x on average (N=20)

Indirect interaction:

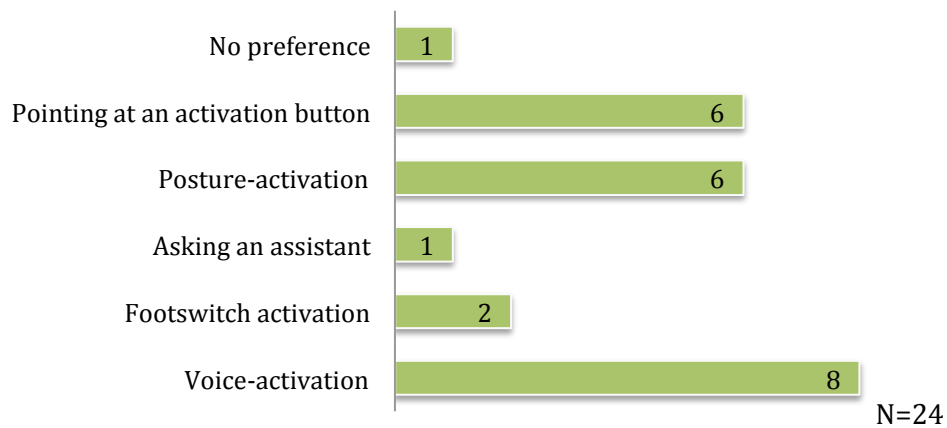
How often do you ask the assistant on average to interact with the patient study on average per surgery?

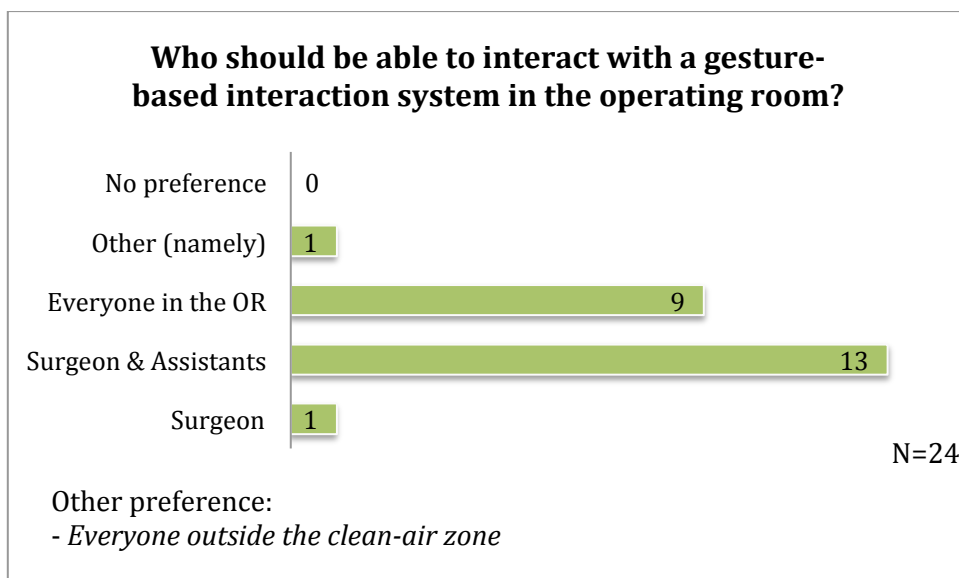
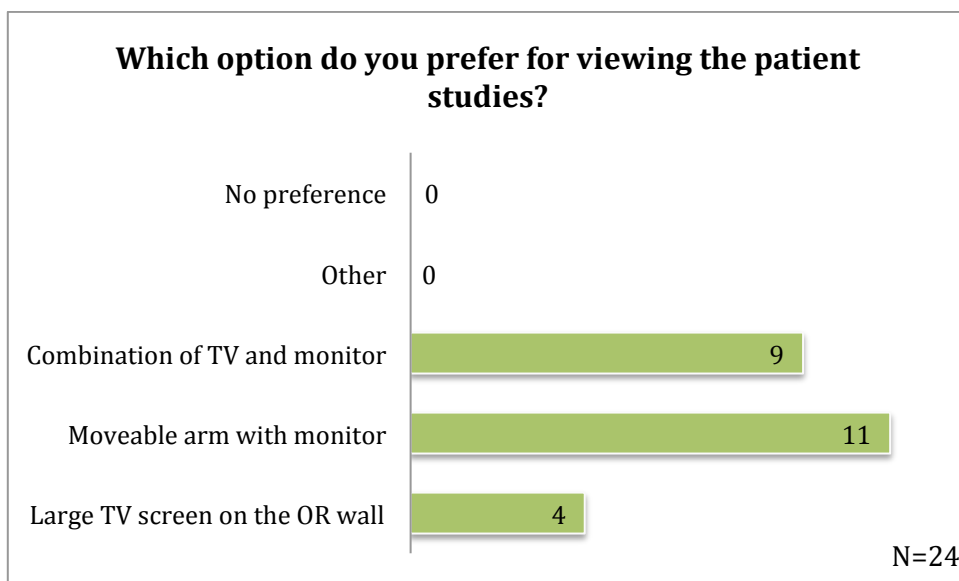
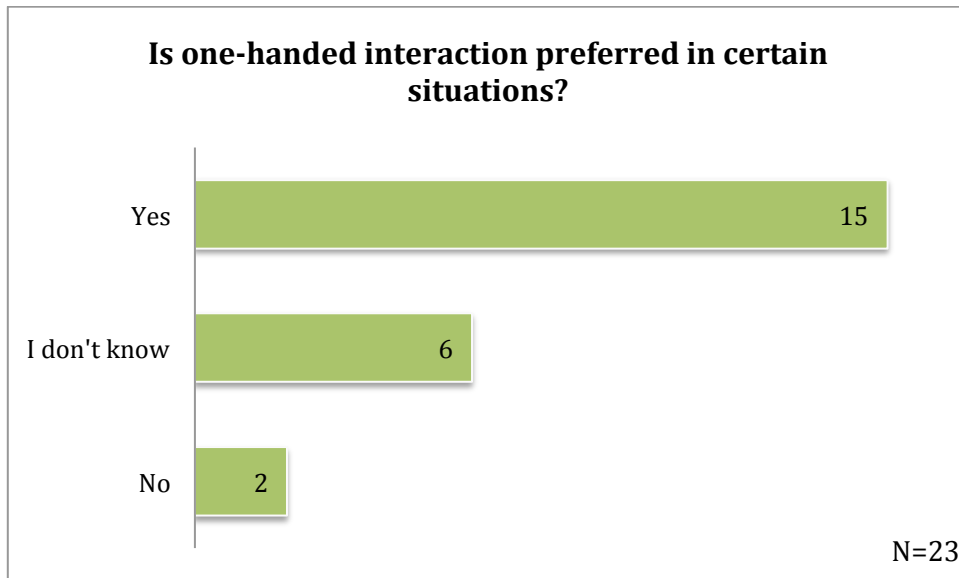
1.51x on average (N=18)

Would you want to interact with patient studies more often (than in the current situation) in an ideal situation where the workflow is not disturbed?

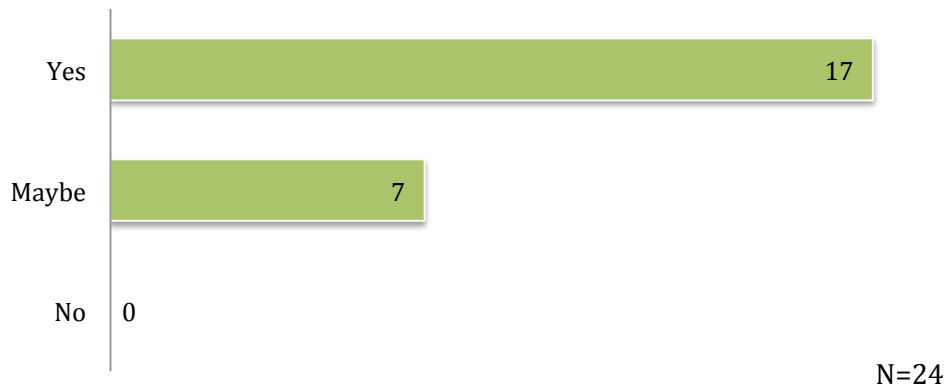


Which (de)activation method of the gesture-based interaction system would you prefer?





Do you regard a gesture-based interaction technique as a promising alternative for the current situation?



Do you have any remaining remarks or questions on this survey or gesture-based interaction methods in the operating room?

- "I don't know if large gestures are convenient in the OR (due to asepsis etc.). I wouldn't mind to use a sterilized wireless (sealed) mouse"
- "Three-dimensional image manipulation would be even better"
- "I am afraid that a lot of time will be wasted on a toy that won't be used in the OR after a while (just like voice-control). I think that I usually merely create a mental image of the radiological images and don't look at these images any further"
- "The questions concerning the amount of times that direct or indirect interaction occur are unclear, is this per surgical procedure? (This doesn't always occur)"
- "Gesture controlled interactions can interrupt the laminar airflow, which is not always desirable"
- "Nice idea! This should be pursued!"
- "Hand and arm movements should be kept small to ensure asepsis, nice idea"
- "The gestures as shown in the video can only be implemented when they do not endanger asepsis"

Appendix B.1

User-based usability evaluation questionnaire overview

Personal data

- 1) Age years
- 2) Gender ☐ male ☐ female
- 3) Handedness ☐ right ☐ left ☐ both
- 4) Specialism within the hospital

How familiar are you with ..

- 5) .. gesture-based technology like the Microsoft Kinect?

Unfamiliar			Very familiar (I own one)			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	2	3	4	5	6	7

- 6) .. touch-screen experience on smartphones (Iphone etc.)?

Unfamiliar			Very familiar (I own one)			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	2	3	4	5	6	7

- 7) .. touch-screen experience on tablet computers/touch pc's?

Unfamiliar			Very familiar (I own one)			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	2	3	4	5	6	7

- 8) .. Laptop/Mac touchpad gestures?

Unfamiliar			Very familiar			
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	2	3	4	5	6	7

- 9) Do you have any other gesture experiences (examples)?

General system usability evaluation

	Strongly disagree							Strongly agree
1) I think that I would like to use this system frequently	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2) I found the system unnecessarily complex	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3) I thought the system was easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4) I think that I would regularly need the support of a technical person to be able to use this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5) I found the various functions in this system to be well implemented	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6) I thought there was too much inconsistency in this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7) I would imagine that most people would learn to use this system very quickly	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8) I found the system very cumbersome to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
9) I felt very confident using the system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
10) I needed to learn a lot of things before I could get going with this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	6	7	

	Strongly disagree							Strongly agree
Zooming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Translation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Rotation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Animating / scrolling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Changing window-level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Changing window-width	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Pointing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Selecting / Measuring	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Cutting out ROI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Clear selection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	6	7	

2) The following gesture requires (much) physical effort

	Strongly disagree							Strongly agree
Zooming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Translation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rotation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Animating / scrolling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Changing window-level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Changing window-width	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pointing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Selecting / Measuring	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cutting out ROI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clear selection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5	6	7	

3) I would like to use the following gesture in a real operating room

	Strongly disagree							Strongly agree
Zooming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Translation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rotation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Animating / scrolling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Changing window-level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Changing window-width	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pointing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Selecting / Measuring	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cutting out ROI	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clear selection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5	6	7	

	Strongly disagree							Strongly agree
4) I found the calibration well implemented	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5) I found the activation/deactivation of gesture-functionality to be well implemented	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6) I had no problem using the system without any physical feedback	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7) I found the system-feedback to be sufficient during interaction	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8) I would prefer small one-handed gestures to wide two-handed gestures	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9) I believe that the used gestures are applicable in a real operating room	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10) I would like to use this system rather than ask an assistant to control mouse and keyboard	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11) I believe that a gesture system could eventually be implemented in the operating room	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5	6	7	

Did you miss any functionality?

Positive aspects of a gesture-based system

Negative aspects of a gesture-based system

Appendix C.1

Selection-task program code (C#)

```
namespace pointExperiment
{
    /// <summary>
    /// Interaction logic for MainWindow.xaml
    /// </summary>

    public partial class MainWindow : Window
    {
        //Declare variables
        public Stopwatch myStopWatch = new Stopwatch();
        public Ellipse targetEllipse;
        public Ellipse previousEllipse;
        public Point mousePos;
        public long time;
        public int count = 0;
        public Ellipse[] ellipses = new Ellipse[16];
        public int misses;
        public int numberOfTrials = 15;
        public int trial = 1;

        //Declare Dictionaries that hold the values of interest
        public Dictionary<string, long> dictTime = new Dictionary<string, long>();
        public Dictionary<string, string> dictDistance = new Dictionary<string, string>();
        public Dictionary<string, int> dictMisses = new Dictionary<string, int>();

        //Start GUI
        public MainWindow()
        {
            InitializeComponent();
            this.PreviewKeyDown += new KeyEventHandler(HandleEsc);
        }
    }
}
```

```

//When GUI is loaded start Experiment
private void MainCanvas_Loaded(object sender, RoutedEventArgs e)
{
    Experiment();
}

//Assign GUI Ellipses to array of Ellipses and set starting target ellipse
private void Experiment()
{
    ellipses[0] = Ellipse1;
    ellipses[1] = Ellipse9;
    ellipses[2] = Ellipse2;
    ellipses[3] = Ellipse10;
    ellipses[4] = Ellipse3;
    ellipses[5] = Ellipse11;
    ellipses[6] = Ellipse4;
    ellipses[7] = Ellipse12;
    ellipses[8] = Ellipse5;
    ellipses[9] = Ellipse13;
    ellipses[10] = Ellipse6;
    ellipses[11] = Ellipse14;
    ellipses[12] = Ellipse7;
    ellipses[13] = Ellipse15;
    ellipses[14] = Ellipse8;
    ellipses[15] = Ellipse16;

    targetEllipse = ellipses[count];
    targetEllipse.Fill = Brushes.Blue;
    TextBlock1.Text = trial.ToString();
}

//Close screen when Escape key is pressed
private void HandleEsc(object sender, KeyEventArgs e)
{
    if (e.Key == Key.Escape)
        Close();
}

//Method to catch cursor click events
private void MainGrid_MouseDown(object sender, MouseButtonEventArgs e)
{
    mousePos = e.GetPosition(targetEllipse);
    mousePos.X = mousePos.X - (targetEllipse.ActualWidth / 2);
    mousePos.Y = mousePos.Y - (targetEllipse.ActualHeight / 2);
    double distance = Math.Sqrt(Math.Pow(mousePos.X, 2) + Math.Pow(mousePos.Y, 2));

    if (distance > targetEllipse.ActualWidth / 2)
    {
        misses++;
    }
    else
    {
        if (count == 0)
        {
            count++;
            previousEllipse = targetEllipse;
            targetEllipse = ellipses[count];
            previousEllipse.Fill = Brushes.White;
            targetEllipse.Fill = Brushes.Red;
            misses = 0;
            myStopWatch.Start();
        }
    }
}

```



```

else if (count < 15)
{
    myStopWatch.Stop();
    time = myStopWatch.ElapsedMilliseconds;
    string key = (trial.ToString()) + ", " + (targetEllipse.Name.ToString());
    dictTime.Add(key, time);
    dictDistance.Add(key, distance.ToString());
    dictMisses.Add(key, misses);
    count++;
    previousEllipse = targetEllipse;
    targetEllipse = ellipses[count];
    previousEllipse.Fill = Brushes.White;
    targetEllipse.Fill = Brushes.Red;
    misses = 0;
    myStopWatch = Stopwatch.StartNew();
}

else if (count == 15)
{
    time = myStopWatch.ElapsedMilliseconds;
    string key = (trial.ToString()) + ", " + (targetEllipse.Name.ToString());
    dictTime.Add(key, time);
    dictDistance.Add(key, distance.ToString());
    dictMisses.Add(key, misses);
    myStopWatch.Reset();
    string trialNumber = trial.ToString();

    if (trial < numberOfTrials)
    {
        trial++;
        TextBlock1.Text = trial.ToString();
        count = 0;
        previousEllipse = targetEllipse;
        targetEllipse = ellipses[count];
        previousEllipse.Fill = Brushes.White;
        targetEllipse.Fill = Brushes.Blue;
    }
    else
    {
        foreach (KeyValuePair<string, long> kvp in dictTime)
        {
            File.AppendAllText("dictionary_time.txt", string.Format("{0}, {1} {2}", kvp.Key, kvp.Value,
                Environment.NewLine));
        }
        foreach (KeyValuePair<string, string> kvp in dictDistance)
        {
            File.AppendAllText("dictionary_dist.txt", string.Format("{0}, {1} {2}", kvp.Key, kvp.Value,
                Environment.NewLine));
        }
        foreach (KeyValuePair<string, int> kvp in dictMisses)
        {
            File.AppendAllText("dictionary_misses.txt", string.Format("{0}, {1} {2}", kvp.Key,
                kvp.Value, Environment.NewLine));
        }
        Close();
    }
}
}
}
}

```

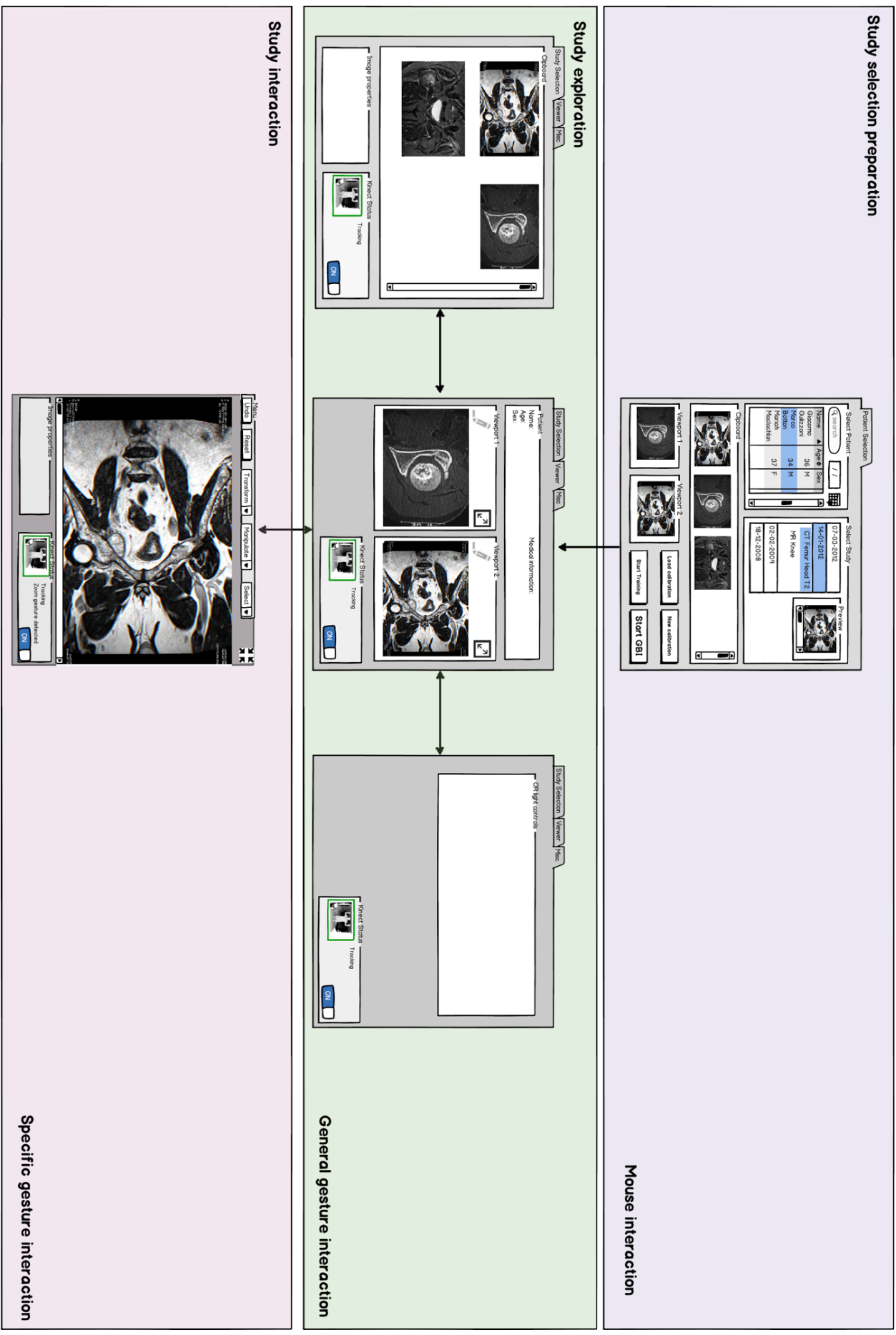
```

}

```

Appendix D.1

Proposed graphical user interface



Appendix D.2

Target selection experiment usability evaluation questionnaire

General participant information:

Personal data

- 1) Age years
- 2) Gender ☐ male ☐ female
- 3) Handedness ☐ right ☐ left ☐ both
- 4) How many hours a day do you work with computers on average

How familiar are you with ..

- 5) .. gesture-based technology like the Microsoft Kinect?

Unfamiliar							Very familiar (I own one)	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
1	2	3	4	5	6	7		

Technique assessment questionnaire

	Strongly disagree							Strongly agree
1) The operation of the cursor was smooth	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2) The mental effort required to operate the cursor was high	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3) The physical effort required to operate the cursor was high	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4) Accurate pointing was easy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5) Target selection was easy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6) Arm fatigue was high during interaction	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7) Overall, the technique was easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8) I would like to use this technique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	6	7	