An Investigation of Multi-modal Gaze-supported Zoom and Pan Interactions in Ultrasound Machines

by

Yasmin Halwani

B.Sc., Qatar University, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2017

Abstract

We are investigating the potential and the challenges of integrating eye gaze tracking support into the interface of ultrasound machines used for routine diagnostic scans by sonographers. In this thesis, we follow a user-centred approach by first conducting a field study to understand the context of the end user. As a starting point to a gaze-supported interface, we focus on the zoom functions of ultrasound machines. We study gaze-supported approaches for the often-used zoom function in ultrasound machines and present two alternatives, One-step Zoom (OZ) and Multi-step Zoom (MZ). A state-based analysis on the zoom functions in ultrasound machines is presented followed by a state-based representation of the gaze-supported alternatives. The gaze-supported state representation extends the manual-based interaction by implicitly integrating gaze input to OZ and offering a gaze-supported alternative to moving the zoom box in MZ. Evaluations of the proposed interactions through a series of user studies, seventeen non-sonographers and ten sonographers, suggest an increased cognitive demand and time on task compared to the conventional manual-based interaction. However, participants also reported an increased focus on main tasks using the gaze-supported alternative, which could offer benefit to novice users. They also report a lowered physical interaction as the gaze input replaces some functions of the manual input.

Lay Summary

This work describes designing an interface for ultrasound machines with an integrated eye gaze tracker. A reported 91% of sonographers experience work-related muskuloskeletal disorders. By delegating some of the tasks performed with the manual inputs of the machine to the eye gaze, the amount of frequent physical repetitiveness needed to perform sonography tasks can be reduced. As a starting point, we target the zoom function in ultrasound machines and investigate approaches to perform zooming with a combination of eye gaze input and manual inputs. We present a field study to understand the context of our target users, followed by an analysis of the zoom functions in the existing ultrasound machines. Our results from a user study performed with sonographers show that the reduction of physical demand, using eye gaze as an additional input, increases both the mental demand and the time on task.

Preface

I was the lead investigator of the work described in this thesis under the supervision of Dr. Fels and Dr. Salcudean. My supervisors presented me with the idea of the project of a gaze-supported interface for ultrasound machines and through field studies and further discussions, the project was narrowed down to investigating gaze-supported interactions for ultrasound machines for routine diagnostic ultrasound exams.

The work presented in this thesis has all been designed and implemented by myself at the Robotics and Control Laboratory and the Human Communication Technologies Laboratory at the University of British Columbia.

External tools include the python wrapper for Ulterius [59], the ultrasound machine communication tool, that was developed by an earlier student at RCL, Samuel Tatasurya and the EMDAT eye gaze analysis tool that was developed at the Intelligent User Interfaces group at UBC [37].

The controls box described in Chapter 4 was implemented by Leo Metcalf, a co-op student at RCL.

The work presented in Appendix D has been published in the Late-Breaking Work category at CHI 2017 [25] and has been presented at the conference in Denver, Colorado, USA as a poster with the latest proposed interaction design described in Chapter 4. The same poster was later presented at the Arab Women in Computing conference 2017 held in Beirut, Lebanon.

The early stages of this work have been presented as a poster and won the third place in the category of student poster presentations at the Qatar Foundation Annual Research Conference in Doha, Qatar in 2016 and published in the Qatar Foundation Annual Research Conference Proceedings [26]. It also won the best student poster award at the annual HCI@UBC forum in April, 2016.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

API	Application Program Interface
BMI	Body-Mass Index
CBD	Common Bile Duct
CHD	Common Hepatic Duct
CW	Clockwise
CCW	Counter-clockwise
DICOM	Digital Imaging and Communication in Medicine
HCI	Human-Computer Interaction
MI	Mechanical Index
MZ	Multi-step Zoom
MZE	Multi-step Zoom Experiment
OOR	Out of Range
OZ	One-step Zoom
OZE	One-step Zoom Experiment
PACS	Picture Archiving and Communication System
PB	Push Button
POG	Point of Gaze
ROI	Region of Interest
SUS	System Usability Scale
TCP/IP	Transmission Control Protocol and the Internet Protocol
TGC	Time-Gain Compensation
US	Ultrasound
UX	User Experience
WRMSD	Work-related Musculoskeletal Disorders

Acknowledgements

I would like to thank Dr. Sid Fels and Dr. Tim Salcudean for their supervision and continuous support during my master's degree. It has been an enriching experience to receive feedback from two professors to direct and fine-tune my work. The time and resources they dedicated for discussing with me all the challenges encountered in this thesis through their busy schedules is very appreciated.

Also, I would like to thank Vickie Lessoway for her enormous support to this project including her help in contacting hospitals, contacting British Columbia Ultrasonographers' Society, recruiting sonographers for user studies, testing a multitude of prototypes and advising on the design of two user studies with sonographers. I extend my thanks as well to BC Women's Hospital and Richmond Hospital for facilitating my observations at their medical imaging departments and to BCUS for distributing the surveys and announcements of participants recruitments for the user studies.

I would like to thank Qatar Research Leadership Program, a member of Qatar Foundation, for their sponsorship of my graduate studies and for their financial support to all other academic-related costs including my travel to conferences and annual QRLP meetings. Special thanks for extending my funding to cover the extra time needed to finish my thesis work.

I would like to thank all the members of both of my labs, the Robotics and Control Laboratory and the Human Communication Technologies Laboratory, for the great support and feedback on my work and for always being available and interested to test new prototypes! Special thanks to Irene Tong for being a great eye gaze tracking buddy. Also, many thanks to Leo Metcalf for his help in designing and implementing the ultrasound controls box and to Dereck Toker from the Laboratory for Computational Intelligence for his help in eye gaze data analysis towards the end of this thesis.

I would like to thank my influencing professors from Computer Science, Dr. Joanna McGrenere and Dr. Karon MacLean, for their inspiring courses in Human-Computer Interaction and User Interface Design. Their engaging lectures and coursework involving interesting and creative projects contributed in shaping my interest in the field and building a strong foundation in Human-Computer Interaction.

Endless thanks go to my community and friends at St. John's College for being the best source of moral support I could ask for and for forming my home away from home.

My deepest thanks go to my loving parents, Aladdin Halwani and Ahlam Karzon, and supportive brothers, Waseem and Hussam, for their patience and prayers and for their sweet daily calls and messages that kept me going throughout the difficult times of my study abroad.

Chapter 1

Introduction

Among the number of tasks involved in image editing-related applications, zooming and panning is one of the basic and most important tasks performed by a user. It belongs to a larger set of functions that require localizing a certain point of interest before performing any further image adjustments. Traditionally, the localization is achieved by the user through variations of mouse and keyboard input, requiring the user to move a pointer on the graphical user interface to the area of interest within the image. Recent advances in human-computer interaction research investigated different input modalities to interact with image editing functions that require localization of an area of interest [56], such as, but not limited to, hand gestures [13], foot pedals [39], or a multi-modal combination of some of these input modalities followed by manual buttons and switches for position confirmation [55].

Eye gaze tracking is one of the recently explored input modalities. Although there are no reported results on an improved effectiveness of gaze-based interaction in comparison to conventional mouse-based input in terms of fine accuracy, many studies show that multi-modal gaze-based interaction has potential in terms of an enhanced speed and user satisfaction [54], given the interface is designed carefully and the user is sufficiently familiar with gaze-based interfaces. Based on early investigations of eye tracking interfaces, Zhai *et al.* argues that, assuming eye gaze can be tracked and effectively used, no other input method can act as quickly as the eye gaze [68].

Eye gaze-based interaction has been investigated both as a stand-alone input [51] and as a multi-modal approach [54] to achieve image-related tasks. The application space for such interaction spans areas ranging from graphic design to medical images inspection. Nevertheless, the majority of those studies investigate the interaction from an abstract point of view and base the interaction on general image editing or image inspection tasks, which is assumed to fit all areas of applications with the same performance level.

In our work, we apply these interactions to the case of ultrasound image acquisition and inspection in diagnostic sonography. As detailed in the next section, differences exist between our work and previous work on multi-modal gaze-supported zooming, such as the types of images used and the added bimanual interaction, which may either amplify or degrade the effectiveness of multi-modal gaze-based interfaces within the context of sonography.

1.1 The Challenges of Sonography

A sonographer is a medical professional who possesses an in-depth understanding of the anatomy, pathophysiology and principles of ultrasound physics to produce medical ultrasound images. A sonographer also communicates with patients while scanning to explain the procedure of an ultrasound exam and the relation between the symptoms and the sonographic image. Prior to an exam, a sonographer has a record of the patient's medical information to be correlated with the resulting ultrasound images and discussed with the physician [24].

Sonographers spend hours of work daily acquiring and modifying parameters of images to be later sent to a physician for further review and diagnosis. As every minute counts toward the throughput of ultrasound exams per day and the health care quality, ultrasound machine interfaces are designed with efficiency in mind to make the access to ultrasound functions as fast as possible.

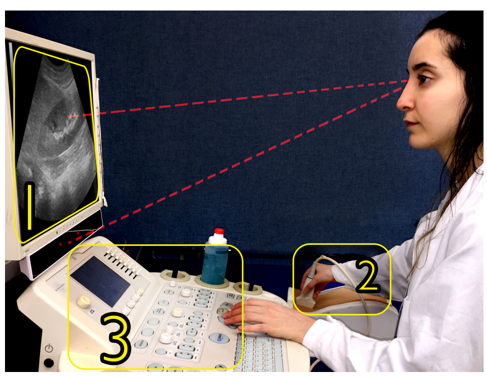


Figure 1.1: The Setting of a Diagnostic Sonographer's Environment and the Sonographer's Three Contexts of Attention: the Ultrasound Image, the Patient and the Machine Controls.

In addition to efficiency, Zhai *et al.* [68] argues that enhancing interfaces with eye trackers has the potential to also reduce repetitive stress injuries for computer users. Figure 1.1 shows the typical environment of a sonographer using an ultrasound machine, where an intricate bimanual interaction is taking place: the dominant hand maneuvers the probe so that the image has the area of interest on the screen while exerting pressure for better image quality, and the other hand manages the details of the acquired image by repetitively reaching for and pressing buttons on the machine controls to change the image parameters and other ultrasound machine-related functions through various knobs, buttons, sliders and occasionally soft buttons on a touch screen.

A study on the prevalence of musculoskeletal disorders among British Columbia sonographers [50] found that 91% of sonographers experience occupational injuries and disorders due to awkward postures, forceful actions and repetitiveness. Furthermore, a survey conducted by our research team revealed that nearly half of the sonographer respondents (N = 48) reported repetitive movements due to menu selection and physical keys interaction as a major cause of their experienced occupational musculoskeletal injuries.

1.2 Pan and Zoom and Ultrasound Machines

Despite the rest of the common drawbacks of gaze-based interaction, such as the Midas-touch problem (the inability of an interface to distinguish between the user's intention of simply looking at an interface element and activating the interface element's function), noisy data and consequent inaccuracies within small areas on the screen, the potential advantages of introducing multi-modal gaze-supported interaction are worth exploring. Reducing repetitive keystrokes and achieving a more efficient performance for imaging tasks serve as our primary motivation to introduce a multi-modal gaze-supported interface to ultrasound machines users.

We select the zoom and pan function as our starting point to a multi-modal gaze-supported interface as it requires a 2-dimensional input, which maps to the type of output an eye tracker device provides: a 2-dimensional point of gaze. Thus, integration is straightforward. In our work, we adopt and modify previously-investigated multi-modal gaze-based approaches for panning and zooming to best fit the workflow of a sonographer performing a clinical ultrasound exam. This particular user scenario is unique due to two main factors:

1. **Types of images** Previous studies primarily focused on images with a substantial amount of content, such as maps [2][54], chip design [6], and multimedia retrieval systems [56], where there are multiple

targets to be acquired and zoomed into across the overall image. In the case of medical ultrasound images, sonographers assess the acquired image from a holistic perspective and, in most cases, only one object of interest is present at a time, such as a gallbladder surrounded by other organs, a tumour surrounded by tissue, or a fetal heart surrounded by fetal organs. In this application area, the purpose of performing pan and zoom is to obtain a higher resolution image of the area of interest, and not to locate a particular target in a dense image as this step is already achieved by moving the probe to the required position over the patient's body.

2. **Bimanual interaction** The application of bimanual interaction as the user interacts with a multi-modal gaze-based interface has the potential to result in a higher cognitive load, which has not been investigated in simpler scenarios explored in earlier work done in this area.

1.3 Contribution

In this work, we explore pan and zoom, particular ultrasound machine functions that are also common in other image editing-related fields. The specific contribution of this work includes:

1. Results from a field study identifying the challenges of sonography, potential advantages and risks of integrating an eye gaze tracker.

2. An analysis of the magnification-related functions in ultrasound machines: types, frequency of use and applications.

3. A state-based analysis of the manual-based and the proposed gaze-supported zoom interactions for ultrasound machines, resulting in a refinement of earlier investigated interactions in the field.

4. Through user studies, we established that:

   (a) gaze-supported zoom increases task time and adds extra mental effort compared to the manual-based alternatives and

   (b) gaze-supported OZ technique is faster compared to the gaze-supported MZ technique on the pan/zoom tasks.

Through the experiments, we measured metrics related to time on task and cognitive load. The results show that, depending on the target and the zoom technique used, gaze-supported interaction differs from traditional manual-based interaction in terms of time on task. In addition, one of the gaze-supported zoom alternatives explored, One-step Zoom, performed better in terms of time on task compared to the other gaze-supported zoom alternative explored, Multi-step Zoom.

By measuring the cognitive load through eye gaze metrics and qualitative evaluation, we find that, although the gaze-supported alternatives introduced are designed to reduce the physical demand of the required tasks, they show signs of an increased mental demand, which, in the case of this design, is due to two main reasons: the novelty of the interface, requiring a higher effort to learn, and the nature of the gaze-supported interaction that requires the user to intentionally gaze at the target area to be magnified while performing the zoom function.

In this thesis, we introduce our work in Chapter 1 and explore earlier work and present the necessary background in Chapter 2. In Chapter 3, we present our field study results including interviews, observations, and a survey distributed to sonographers. In Chapter 4, we present the proposed gaze-supported zoom interaction design. We test the proposed design in two user studies: the first is a context-free game-based user study presented in Chapter 5 and the second is a context-focused clinical user study presented in Chapter 6. We present our final conclusions and recommendations for future work in Chapter 7. Appendices include supplementary material related to the content presented throughout this thesis.

Chapter 2

Background and Related Work

In this chapter, we introduce the necessary background needed for understanding our work in the following chapters. We discuss the research approach we follow, eye gaze tracking technology and relevant work in the field, types of ultrasound machines, and we analyze the input devices and software design of the target ultrasound functions in the target ultrasound machines class.

2.1 Eye Gaze Trackers as Input Devices

2.1.1 An Overview of Eye Gaze Metrics and Eye Gaze Tracking Technology

Eye gaze tracking is the technology of tracking eye-related measures, including the position of the point of gaze, the size of the pupil, and other characteristics related to eye movement. There is a multitude of eye tracker types that perform the same tasks with different levels of accuracy and obtrusion. Vision-based tracking grew to be the most common method of eye tracking in the recent years due to its unobtrusiveness and ease of integration with existing user interfaces. By projecting infrared light in the direction of the user's eyes and recording the reflection with an infrared sensor, the eye tracker compares the corneal reflection to the pupil position to determine the relative target the user's point of gaze is positioned at.

In order to make sense of the eye tracker's data, and consequently use it as an input channel in user interfaces, one must first understand the basics of human eye behaviour. The human vision spans about 180 degrees, with the highest visual acuity around the fovea, which spans only about 2 degrees. In the context of eye tracking applications, two eye movements are of most importance: saccades and fixations. In the absence of a moving stimulus, the eyes jump rapidly from one area to another about three to four times per second. This type of sudden ballistic eye movement is called a saccade, which is very jittery in nature. Another type of eye movement is the fixation, which occurs when the eye is focusing on one particular spot. On average, fixations are still jittery and last between one-tenth to one-half of a second [8].

Current technology is able to capture the foveal vision's saccades and fixations with varying accuracies depending on the hardware specs and software algorithms used for detection. Bojko [8] lists the hardware specifications that control the quality of the output of the eye tracker: sampling rate, accuracy, precision, head box size, monocular/binocular tracking and pupil illumination method. Currently available trackers perform at a sampling rate of 25 Hz to 2000 Hz, depending on the price range. Although basic eye trackers with low sampling rates sufficiently identify areas of fixations, higher sampling rates directly impact the accuracy of measuring the lengths of those fixations. The accuracy of the eye tracker is the deviation between the recorded point of gaze and the actual point of gaze of the user [8]. If calibrated well, eye trackers are able to record the user's point of gaze at an accuracy of 0.5 to 1 degrees of visual angle. The exact number of pixels for a specific degree of visual angle is dependent on several factors: the distance between the user and the eye tracker, the resolution of the screen and the dimensions of the monitor, as detailed in Appendix A. Precision measures the tracker's ability of reproducing successive identical points of gaze. Head box size defines the flexible box area (width, depth and height) around the user's head where the user is able to move freely without leaving the eye tracker's field of view. Currently available eye trackers are mostly binocular, which means they track both eyes of the user. Although data can be sufficient from one eye, averaging the data from both eyes yields higher accuracy and precision and can also provide gaze depth. There are two methods by which pupils are detected, which serve to create contrast between the eye and the pupil to allow for better detection: bright-pupil and dark-pupil methods. Each method has its merits and is effective under different circumstances based on the environment's brightness and the physical characteristics of the user's eyes. Most eye trackers switch automatically between methods to optimize for detection in varying conditions.

In addition to hardware capabilities of eye trackers, many commercial eye trackers are provided with developer API's offering methods that perform most of the needed post-processing of eye gaze data, including filtering for saccades and fixations, extraction of eye depth (3D) data and varying calibration options. Such products are ideal for human-computer interaction-centred research as they allow the researcher to focus time and effort on the usability of eye tracking rather than on the technical computer vision and data processing aspects.

Due to differences between users in the way they look and behave, individual calibrations of users are required prior to using an eye tracker [8]. A calibration procedure typically includes a set of targets that the user has to sequentially fixate his or her eye gaze on in the order required by the system. Internally, the eye tracker maps each of these targets to the appearance of the eye and performs interpolations to the rest of the visual field to interpret the rest of the eye movements. Therefore, the more the calibration targets, the more accurate the eye tracking. Most eye trackers allow the developer/researcher to set the number and position of calibration targets. The conditions under which the users have calibrated should remain the same for the rest of the use of the eye tracker; this includes the level of ambient lighting in the room and the relative position between the user and the eye tracker. If any of these conditions change, a re-calibration is required. Thus, the amount of re-calibration needed is dependent on the individual user behaviour.

In addition to technical aspects, human aspects also contribute in the quality of eye tracking. As mentioned, accurate calibration is a primary pre-requisite to accurate eye tracking. Additionally, the type of eye wear a user has and the user's age influence the tracker's accuracy.

2.1.2 Eye Gaze-supported Interfaces

Eye gaze trackers can be used either as active input devices or as passive monitoring devices. The eye gaze tracker can be used passively to evaluate interfaces in terms of identifying the areas of attention the users' eye gaze fixates at while interacting with a newly developed interface being evaluated by designers. In terms of active input, the most common application area is developing input mechanisms for users with disabilities, especially for eye typing [27]. However, more applications are starting to emerge, especially when paired with other input mechanisms, such as in gaming and entertainment [57] and recently in mobile devices [9]. Another area where eye gaze tracking-supported interaction is investigated is in facilitating software development environments, such as the work presented in [22], which presents an eye gaze-supported system aimed to enhance source code navigation by enabling the user to activate actions by looking at certain triggers. Another area is the combination of eye gaze input with with other forms of modality, such the work presented by Chatterjee *et al.* [12] investigating multi-modal gaze and gesture recognition.

Eye gaze tracking is explored in the area of medical device interfaces, such as the work presented by Tong *et al.* [61] that investigates gaze-based interaction in the field of robotic surgery. Another application presented by Tan *et al.* [58] introduces a dynamic control-to-display gain mouse movement method with the control of an eye gaze input to facilitate target prediction. In their work, mouse movement is reduced by up to 15% for medical image analysis tasks.

In addition to position-related measurements, eye-tracking technology has the ability to measure other eye parameters, such as pupil dilation, which reflects the level of cognitive load the user is experiencing. Studies such as [48] investigates several psychophysiological measurements for task load analysis and finds pupil dilation to be one of the most responsive measures that can reveal cognitive load in real time. Therefore, eye gaze tracking can also be used to design adaptive user interfaces as the user's cognitive load is being continuously assessed.

2.1.3 Eye Gaze-supported Zooming

Based on early research by Zhai *et al.* [68], users look at the target before they initiate the action to acquire it. In the case of mouse-based zoom interactions, a user first looks at the target to be zoomed into, moves the cursor to indicate the position of interest then confirms the zoom action with a mouse click or a button. Integrating eye gaze tracking with zoom interactions will eliminate the need for the step of moving a cursor towards the target. The user simply has to look at the position of interest then confirm with another input modality to zoom.

Using eye gaze input to aide in zooming is investigated in prior work. As an example, Mollenbach *et al.* [46] performs a comparison of pan and zoom between two modes of input: mouse-activated and gaze-activated. They also study the type of navigation technique that is best suited for gaze-based interaction. The first technique is a search task where the user has to locate a small target within a large collection of shapes, and the second technique is a localization task where the user is required to zoom into a predefined sequence of clear targets. In the results, the authors argue that gaze as a navigational input can be very effective if used for the right type of task. Their experiment yields an improvement of 16% of task performance with the target selection task, which matches the type of zoom interaction performed on ultrasound images by sonographers. In ultrasound imaging, once the probe has been placed at the correct position, the visual target search task is not as intensive as the one presented in [46], as ultrasound targets take larger portions of the screen, therefore the visual task performed is more of a target selection task than a target search task. Similar to other work [2], Mollenbach *et al.* [46] also use "edge-scrolling", with the speed of the scroll movement proportional to the vector between the centre of the image and the point of gaze.

Stellmach *et al.* [56] presents a focus-plus-context interaction based on a gaze-supported zoomable interfaces, which investigates the interaction in combination with keyboard input and with a touch-and-tilt-sensitive hand-held device. An advantage that multi-modal gaze-based has over gaze-only interaction is the avoidance of both the Midas Touch problem and the delay caused by dwell-time activation. In their formative user study, they additionally find that multi-modal gaze-based interaction supports multi-tasking that is otherwise not achievable with gaze-only interaction, such as simultaneously zooming and panning.

Another study presented by Stellmach *et al.* [54] is an application of simple pan and zoom through multi-modal gaze-based interaction for Google maps. Similar to the approach taken in [56], the authors compare different modalities of interaction integrated with eye tracking, including a mouse scroll wheel, the orientation of a hand-held device and touch gestures with a hand-held device. Typically, traditional mouse interaction without the integration of eye gaze yields the best results in terms of time, accuracy, spatial awareness, ease of learning and intuitiveness, due to the user's familiarity with mouse-based systems. However, multi-modal gaze-based pan and zoom integrated with mouse input follows closely to the pure mouse interaction in the aspect of perceived speed and spatial awareness.

The work presented in [46], [56] and [54] is different from our work as the types of images investigated are different in nature to ultrasound images. Nevertheless, it is promising to see that mouse interaction integrated with eye gaze for simple pan and zoom tasks performs better in terms of speed than the rest of the multi-modal systems tested.

Similar to [54], the work presented in [2] investigates the user's performance of zooming and panning in Google maps through four modes: gaze and dwell, mouse only, mouse and gaze, and head movement and eye gaze. In our work, we adopt the DTZ (Dual-to-Zoom) approach, which combines eye gaze with mouse clicks. In [2], the authors use the user's gaze input to localize the area of interest and the mouse right and left clicks for zooming in and out. Moreover, they define pan regions, where the zoomed image pans if the user selects a pan region. The results of this paper suggest that using multi-modal gaze-based control with mouse for zooming in and out is the best after the traditional mouse control in all aspects including time and accuracy. Although stare-to-zoom performs similarly to dual-to-zoom in some respects, staring at images will hinder the speed at which the scanning task is carried out and will increase eye fatigue for prolonged use. Similar to some other work in gaze-supported interaction, the pure manual interaction outperforms the gaze-supported interaction in all quantitative measures.

One of the influential papers to our system design is the work presented by Kumar *et al.* in [42]. In the beginning of the paper, the authors acknowledge that the system they intend to design will not outperform the default manual interaction (in the case of their work, the regular mouse and keyboard), and they stress on the fact that it is not intended to *"replace or beat the mouse"*. Their work is aimed at designing an efficient gaze-supported interaction for those users who opt not to use a mouse depending on their abilities. In their results, users show higher preference for the gaze-supported alternative in terms of speed, ease of use and user preference. However, in terms of accuracy, the mouse alternative is preferred. This work also finds that, with the recruitment of 20 participants, the eye gaze tracker works better for some participants, depending on the posture and the calibration quality during the experiment. It also depends on the individual participant behaviour, for instance, they report on one subject that squinted and laughed a lot during the experiment, which hinders the quality of eye gaze tracking.

In earlier work that combined eye gaze with foot pedals and mouse input [39], results show that participants are able to beat the time on task using the multi-modal gaze-supported interaction compared to the traditional manual-based interaction. However, this difference is not statistically significant. Similar to [42], prior to collecting results from their work, the authors expect the novel interaction technique to not outperform the conventional manual-based interaction, but expect it to be at least comparable with an improved user experience.

One of the studies that shows that eye gaze interaction can actually outperform manual interaction is the work done by Fono and Vertegaal [17]. In their work, they claim that their gaze-supported system achieves an improvement of 72% over the typical mouse interaction. However, the task space is quite different from the image optimization and analysis tasks we target for sonography, as they investigate switching and zooming between windows instead. The targets they investigate, windows, are well-defined and make up a large portion of the screen, which makes them easier to select even with jittery input such as eye gaze. In our work, the target is only defined upon image acquisition by the sonographer and is not known a priori to the system. In addition, the size of the target, as discussed in later chapters of this thesis, can vary.

2.1.4 Eye Gaze-supported Panning and Scrolling

Another common application of gaze-supported interaction is automatic scrolling for on-screen reading, such as, but not limited to, the work presented in [53]. Although their results show that there is no significant difference between automatic and manual scrolling, it is worth investigating further as our application area is quite different as we deal with images rather than text. In our work, we adopt similar approaches to pan images by detecting when the user looks at the edges of the image, instead of the edges of text limits, and move the image accordingly. This is similar to the gaze-supported panning approach used by Adams *et al.* [2] and Mollenbach *et al.* [46].

2.2 Ultrasound Machines: Applications and Target Users

Ultrasound imaging is used in a large set of clinical applications. With it, comes a diverse set of users in diverse settings. By taking these factors into consideration, manufacturers of ultrasound machines created designs with a variety of layouts and capabilities to best suit the different types of ultrasound operators and different applications. A common application where ultrasound imaging is used in is routine ultrasound exams in ultrasound rooms. These exams are concerned with producing high quality images of specific targets and performing related measurements. Examples span areas such as abdominal, cardiac, obstetric, gynaecologic, vascular, musculoskeletal and other general ultrasound exams. Another application area where ultrasound imaging is used is in point-of-care, where urgent ultrasound exams are required in cases of emergencies for diagnosis. Ultrasound imaging is also important to aide physicians who perform interventional procedures to guide their primary tasks such as needle-insertion. Similarly, surgeons use ultrasound imaging to help guide them during their operations and show the underlying anatomical structures. The user-machine interaction in each of these application areas differs from one another, which drive the design of a variety of ultrasound machine types including, but not limited to, platform-based machines and portable machines, as shown in Figure 2.1.

2.2.1 Machines Targeted for Routine Ultrasound Exams

The ultrasound machine operators in this application area are typically sonographers: specialized medical professionals trained with an extensive knowl-

(a) GE Logic E9: a Platform-based Ultrasound Machine Typically Used in Routine Ultrasound Exams. (Image Source: www.kpiultrasound.com)

(b) Sonosite X-Porte: an Example of an Ultrasound Machine with a Completely Touch-based Input Interface. (Image Source: www.sonosite.com)

(c) Clarius: a Hand-held Ultrasound Machine Device. (Image Source: www.clarius.me)

(d) GE NextGen LOGIQ e: a Tablet Ultrasound Machine Suitable for Point-of-care Applications. (Image Source: www3.gehealthcare.com)

Figure 2.1: A variety of ultrasound machine interface designs are available for a variety of target users and applications.

edge on ultrasound image acquisition and other related ultrasound functions. Routine ultrasound exams are performed in ultrasound rooms that are prepared with all the necessary equipment and suitable environment for optimal image acquisition.

During these exams, the sonographer is constantly switching between three contexts of attention: the ultrasound image, the machine controls and the patient, as shown in Figure 1.1. However, the sonographer's main focus is on the production of the acquired ultrasound images and the various accurate measurements performed on specific anatomical targets. Additionally, image acquisition in this area requires the sonographer's management of bi-manual input to the machine: ideally, the dominant hand is in charge of the main task, namely, acquiring the image, and the non-dominant hand manages the properties of and any operations, such as measurements, performed on the image.

Given these interaction factors and the background of the intended set of users, machines in this application area support certain layouts, ergonomics, ultrasound functions and imaging capabilities. Typically, they are platform-based machines with a layout providing a large set of manual inputs to facilitate the user's rapid access of the non-dominant hand to the machine's various functions. Given the variety of exams that can be performed with the machine in this application area, ultrasound machines designed for sonographers possess high processing powers, frame rates, image quality and large monitors to aide in producing the best image possible with substantial detail. Machines in this area typically support multiple types of transducers and are equipped with advanced application-specific technologies, such as, but not limited to, special measurement packages and 4D visualization tools. Some machines also support customizable software and automatic image optimization functions to increase productivity. Platform-based ultrasound machines are also designed to be as ergonomic as possible since they are the primary use of sonographers for prolonged hours during their workday.

2.2.2 Machines Targeted for Point-of-care

Unlike the first application area, operators of ultrasound machines in point-of-care may not necessarily be sonographers and can have varying clinical backgrounds. The setting of point-of-care ultrasound imaging takes place outside the ultrasound room, where scanning and diagnosis are done, for example, at an ICU or at an emergency vehicle. While the image acquisition and optimization tasks may have some common similarities to the case of routine exams, the required mobility of the ultrasound machine impacts the interface layout and offered machine capabilities.

Ultrasound machines aimed at this application area are designed with varying levels of compactness, durability, and portability. A large variety of options are available including portable machines, tablet machines and even newer versions of designs including hand-held pocket-sized machines. However, most of these machines are limited in terms of imaging options, capabilities and ergonomic layout due to the higher priority of providing a portable system. For instance, systems such as SonoSite X-Porte and GE Venue 40, as shown in Figure 2.1b, are completely based on touch-screen input with a few soft keys, which do not provide quick access to ultrasound image optimization and patient data control. However, these types of machines are considered excellent options for point-of-care due to their lightweight, long battery life, and layout design, which provides access to the most frequently-used imaging options only, which makes it suitable to users with limited background in sonography, such as physicians.

2.2.3 Machines Targeted to Aide Other Clinical Tasks

In contrast to the two aforementioned applications, physicians use ultrasound machines for a different goal. Ultrasound images used in interventional needle-insertion or in intraoperative guidance are aimed at providing the physician with a visualization of the underlying anatomical structure. In other words, the physician's main attention is directed elsewhere and the ultrasound image serves as a tool to help them perform their primary task. In such applications, the setup and the interaction are quite different from platform-based ultrasound machines used in routine scans. The user's main focus is on the primary task (inserting a needle or performing a surgical procedure) and the ultrasound machine is only used to acquire an image to aide the guidance of the primary task. For interventional procedures, the dominant hand handles the primary task of the insertion of the needle and the non-dominant hand holds the probe to provide the visual context to the primary task. For surgeries, sometimes the ultrasound probe is handled by a second user altogether as the surgeon's hands are occupied with surgical tools, while maintaining a parallel attention on the ultrasound image and the patient.

Similar to point-of-care, machines suitable for this area of application are usually portable and do not offer a wide range of options given the machine operator's limited background and need of ultrasound image optimization. Provided image optimization is not the primary concern of users of such applications, the lack of physical controls and reduced access to image optimization settings is not considered as a limitation but as an advantage as the main focus of the interface is on the image to aid performing other types of tasks such as guide needle insertion or do a quick investigation at the ICU.

The Target Ultrasound Application Area of This Research

In this research, we target the interface design of the class of ultrasound machines used for routine diagnostic ultrasound exams, such as the type of machines shown in Figure 2.1a. This is due to the fact that the primary users of this machine, sonographers, interact with the machine's various functions more frequently and for prolonged periods of time compared to the users of other types of machines. In this application area, the acquisition of ultrasound images is the main goal of the users, making the use of the machine their primary task. Specifically, we target the machine's image magnification-related functions of High-resolution zoom and Low-resolution zoom.

2.3 Ultrasound Machine Interface Design Analysis

Before introducing the new gaze-supported approach, we must first understand the existing design of magnification-related functions in ultrasound machines in terms of input device interaction and software interaction. This will help create an informed gaze-supported interaction design for the pan and zoom function in ultrasound machines.

2.3.1 Input Device Interaction

In this section, we discuss the theory behind the design of input methods related to ultrasound machine controls. Specifically, we target classical HCI theories that define the capabilities of input devices and map it to the types of tasks performed by users of these devices and the target functions we are investigating. We apply these theories to current practices of ultrasound machine controls layout design and discuss the advantages and drawbacks of the different approaches taken in in the various premium and high-end models of ultrasound machines used in routine diagnostic sonography.

Buxton's work on the three-state model of graphical input [11] maps the demands of interactive transactions to the capabilities of input transducers. In his work, he argues that all input transducers exhibit at least one of the three defined states. Therefore, by identifying the tasks to be performed with an interface, a designer will be able to select the most suitable input transducers that will optimally serve those tasks.

The first state is “state 0”, which is described as “out of range” or OOR. In this state, the system is not receiving information from the input device. An example is a 2D stylus: when it is lifted off a tablet, the system is unable to tell the position of the stylus while not in contact with a surface. The second state is “state 1”, which has a different description based on the input device. A consistent abstract description of state 1 across all devices is a trackable continuous signal. For example, a mouse or a joystick is always in state 1 unless another action takes place, such as a button press. State 2 occurs when an additional simultaneous action that supplements state 1 takes place. As mentioned earlier, pressing a button while moving a mouse (such as the mouse’s left button) will transition the mouse’s state from state 1 (tracking) to state 2 (dragging) for manipulating a desktop graphical user interface. Another example for state 2 is a threshold pressure for a pressure-sensitive stylus that will transform its state from state 1 (tracking) to state 2 (dragging, or inking, if the stylus is used for a drawing application, for instance).

Although simple in theory, the three-state model can support the representation of more complicated systems. First, multiple states of the same type can be present in an input device. Buxton brings an example of a mouse as it has two buttons, which could serve two different purposes for a particular application and thus a mouse with two buttons has three states, of which 2 are of the same class (state 2). He also highlights the difference between continuous and binary transactions and how the three-state model represents them. For instance, pointing is a continuous task and clicking is a binary task. In a three-state model diagram, both are represented similarly. The difference is in terms of implementation, as clicking tasks are treated as a state 1-2-1 transactions, without motion within state 2. In other words, as long as the signal value sent by the input device to the system stays constant while in state 2 (the device isn’t moved), the transaction is classified as a clicking task.

Lexical and Pragmatic Characteristics The model presented earlier helps in mapping the tasks to the capabilities of the input devices. These capabilities, however, could be shared across a class of different input devices. For instance, looking at the state characteristics table in Buxton’s work [10], a trackball, mouse and a joystick are all capable of supporting states 1 and

2. Similarly, tablets, touch tablets, touch screens and light pens are capable of supporting all 3 states. Therefore, further interaction theory should be investigated related to the human motor/rotary system and its impact on the choice of input devices made by a layout designer. An earlier work presented by Buxton [11] investigates the lexical and pragmatic considerations of input devices that targets the classification of continuous hand-controlled devices in terms of the property sensed, the number of dimensions and the muscle group involved. Buxton's work is later extended by Kobayashi *et al.* [40], which created a new classification taxonomy that includes auditory and visual devices in addition to hand-controlled devices.

The most important output of Buxton's work on the consideration of lexical and pragmatic characteristics of input devices is the tableau of continuous input devices [10], which we use to further narrow down the mapping of input devices used for the types of transactions required by image magnification-related functions in ultrasound machines. The rows in the tableau classify the input devices based on the property sensed: position, motion and pressure. The columns, on the other hand, classify the devices based on the number of dimensions they control. Furthermore, rows are divided into mechanical versus touch-based control for each type of property sensed. In addition to defining the type of input devices suitable for a particular transaction based on the three characteristics discussed, the tableau is also helpful for finding equivalences and relations between devices.

Input Devices for Bimanual Interaction Provided ultrasound machines are controlled bimanually, the use of the left hand to interact with the physical layout of the ultrasound machine also requires an investigation of the types of input devices optimized for use by the non-dominant hand. We found the series of studies performed by MacKenzie *et al.* [43] and Kabbash *et al.* [35] to be of use for this particular research, as will be detailed in later sections.

2.3.2 Image Browser Interaction

The work presented by Plaisant *et al.* [47] specifies the browser interface design's presentation and operation aspects. They classify the tasks of an image browser as follows: image generation, open-ended exploration, diagnostic tasks, navigation, or monitoring. Sometimes, a combination of one or more of these tasks may be needed, in which case the image browser must be carefully designed to support the appropriate presentation and operation aspects of both (or more) tasks.

Plaisant *et al.* [47] also describe the presentation aspects of an image browser. The presentation of an image browser has static and dynamic aspects, where the first is concerned with the layout of the views presented to the user, and the latter describes the update methods of the presented layout(s). A number of different operations can be performed with an image browser, such as inspecting details, moving between important pre-defined areas within an image, navigation, and more. However, the authors place a higher emphasis on the pan and zoom functions, since they are basic manual operations performed by the user required by most image browsers. Additionally, they categorize a number of operations that they recommend designers to automate in their image browser interfaces to make it easier for users to concentrate on their main tasks.

2.3.3 Magnification-related Functions

Low-resolution Zoom There are two basic variations of zoom functions in ultrasound machines. The first one is the Low-resolution: quickly-accessed by the user, magnifies the entire image from the default centre position in discrete steps within the detail view through an input device, and de-magnifies (zooms out) in the same manner through a reverse action of the same input device. The input device is usually a continuous rotary knob, as in the case with all premium and high-end machines researched and observed, and sometimes input buttons, as in the case of lower-end and some tablet and portable machines. Low-resolution zoom does not perform any image optimization on the magnified image and is used for quick investigation of a particular area of the image.

High-resolution Zoom The second zoom function variation is the High-resolution zoom. As shown in Figure 4.5, High-resolution zoom provides the sonographer with a higher level of control over the magnified image. Moreover, the scan frame rate (also known as the ultrasound temporal resolution) and line number are typically automatically optimized as the imaging sector is narrowed by the machine through magnifying the image, which makes it a preferred zoom alternative when a sonographer is likely to zoom into a part of an image that is rapidly moving and requires a higher temporal resolution image acquisition. Provided the inherent change in image acquisition as the probe limits its field of view to capture the magnified area only, panning within the zoomed image is not technically possible as a post-processing function. An ultrasound probe must have the capability of actively moving its lateral field of view to support panning of the magnified area.

In High-resolution zoom, once the pre-zoom mode is activated, the border of the field of view to be magnified (or the "zoom box") is displayed over the image (either in the overview or the detail view, or both) with a default size and position, which is initially placed at the centre of the image. No magnification takes effect until the zoom action is confirmed with the selected dimensions and position of the field of view. The zoom factor in this case is implicit, as the dimensions of the field of view determine it. Similar to Low-resolution zoom, magnification is restricted by restricting a maximum and a minimum width and height dimensions of the adjustable field of view set by the sonographer. This particular zoom approach has no reverse zoom action. Therefore, zooming out simply resembles restoring back the image to its default magnification outside the selected field of view, which is referred by Plaisant *et al.* [47] as implicit zoom out or "undoing".

Hybrid PanZoom Newer machines support a hybrid of high resolution and Low-resolution zoom. As discussed earlier, some machines name it the PanZoom option. This alternative serves as a more controlled Low-resolution zoom function. Similar to High-resolution zoom, the position and size of field of view to be magnified is controlled manually. However, since it is essentially a post-processing function on the acquired image, a sonographer is able to pan the magnified image within the rest of the overall field of view. This option is used when the temporal resolution of the image is not a main concern for the sonographer, but control over the dimensions of the acquired magnified image is still of importance, as shown in Figure 4.5.

In both low resolution and hybrid PanZoom approaches, panning is performed by first changing the image mode to panning mode, then moving the trackball in the desired pan direction. Since there is no equivalent for it in Plaisant *et al.*'s [47] taxonomy of pan operation, we added pan by "cursor movement" to the taxonomy as the trackball functions as a cursor in ultrasound machine interfaces.

Transducer-based Panning One notable interaction we observed with sonographers is sometimes their tendency to manually pan the area of interest by moving the transducer over the patient instead of using the software-based pan function, even if the target is completely visible within the range of the image acquired by the transducer. Although software-controlled panning is much more stable (as the sonographer might lose the target of interest by slightly moving the probe more than needed), we suspect that panning with the probe is preferred in some cases as it does not require any context switch to perform the software-based pan. By context switch, we mean the extra interaction with the machine controls to select the panning mode and moving the trackball in the desired direction. It is an additional action that, sometimes, could be avoided if the transducer is carefully moved over the patient.

Depth Upon observations of several ultrasound exams and informal discussions with sonographers, as we describe in Chapter 3, we find that the zoom and depth functions are used interchangeably to bring a particular area of interest into the central view of an acquired image. In terms of operation, both functions are identical (compared to the operation of the Low-resolution zoom only). However, sonographers prefer to use the depth function to the zoom function whenever possible, especially if the target is located close to the surface. The maximum allowed depth by the transducer is dependent on the frequency it is set at. The higher the frequency, the lower the penetrable depth and the higher the lateral resolution. More importantly, similar to High-resolution zoom, decreasing the depth improves the temporal resolution as well. In this case, the temporal resolution is improved as it decreases the pulse repetition period. Therefore, unless the target is located at a high depth, sonographers typically prefer to control the depth of the image to bring a particular target into central attention as it acquires images with better lateral and temporal resolutions instead of using the available zoom functions.

2.4 Conclusion

We present a brief background on eye gaze tracking technology and the common applications that use the gaze input for direct control. Previous research in the field of gaze-supported zoom control has been already explored with different results based on the types of images zoomed and the modality of the interaction. Two studies which show that gaze-supported zoom is faster than manual-based zoom is that presented by [46] and [17]. However, the types of images and tasks presented in these studies differ from the target application of zooming into ultrasound images within the context of sonography.

Later, we classify ultrasound machines based on the application area and the target set of users. In this research, we target the zoom interfaces of ultrasound machines used by sonographers for routine diagnostic sonography.

Finally, we present classical HCI theory on input devices based on Buxton's [11] three-state model and devices' lexical and pragmatic characteristics [10] and bimanual interaction [43] [34]. In addition, we provide an overview of image browser design principles based on Plaisant *et al.*'s work [47]. We refer to these theories in Chapter 4 to explain the input device and image browser design of the targeted ultrasound machine's layouts and help create an informed design of gaze-supported zoom interaction.

Chapter 3

Field Study And Observations

In this chapter, we present the methodology and results of the conducted field study. We perform observations, conduct interviews and distribute a survey to understand the environment of our target users: sonographers. We use these results to help us build an informed gaze-supported interaction with the ultrasound machine as we analyze the potential benefits and risks associated with eye gaze input integration.

3.1 Methodology

User-centred design is "*a broad term to describe design processes in which end-users influence how a design takes shape*" [1]. In the field of medical device development, following a user-centred design approach is essential to help support the end user's needs better, since the target user group is small, with unique experiences and requirements. User-centred design involves iterative cycles of the following: defining user objectives, collecting requirements, evaluating design alternatives and testing the proposed system with end users.

Studies on ultrasound machine interface design followed a user-centred design approach for over 20 years. The work such as [5] introduces a number of design alternatives and changes to the ultrasound machine interface based on a user-centred design approach. In addition, work such as presented by Martin *et al.* [45], recommends a list of user-centred research methods to ensure ergonomically-designed medical devices. In our work, we follow some of the methods recommended, such as contextual inquiry, in conjunction with a series of usability tests and heuristics in the early stages with expert sonographers.

In this chapter, we present results from interviews with sonographers, ultrasound scans observations, and a survey distributed to the members of the British Columbia Ultrasonographers' Society. We identify some of the challenges that sonographers face in sonography and conclude with discussing the potential risks and benefits of deploying a gaze-based interaction technique based on observed user behaviour, ultrasound machine capabilities and the clinical environment setting.

3.1.1 Observations

The ultrasound machine, functions, and exam duration all differ based on the specific ultrasound scan being performed and other factors related to the patient's physiology and the sonographer's experience. To get a practical view of these factors and the feasibility of integrating an eye tracking system with the machines given such diversity, various types ultrasound exams were observed, including general, obstetric, breast, renal, vascular, abdominal and echocardiography exams at two different hospitals for two full days summing up to a total of 18 ultrasound scans. In addition, we observed an ultrasound-guided breast biopsy procedure.

3.1.2 Survey

An online survey was conducted to get an in-depth understanding of the observed challenges with sonography in practice. The aim of the survey is to understand the ultrasound machine interaction from a user's perspective to help in directing the design and requirements of the new eye gaze-supported ultrasound machine interface. The questions asked to survey respondents relate to an ultrasound machine user's daily interaction with the machine, musculoskeletal disorders due to work injuries, and emerging ultrasound machine technologies. A total of 66 responses were received, of which 48 are completed responses. All survey questions can be found in Appendix B.

The survey is designed for the distribution to both sonographers and radiologists with experience in ultrasound, including radiologists who perform ultrasound-guided interventional procedures, to understand the differences in human-machine interaction between the two user groups and the interface requirements based on the interaction. However, due to difficulties distributing the survey among radiologists, only sonographers' responses are considered in this analysis. Responses have been actively incoming since the survey was distributed on Feb. 17, 2016 until Apr. 24, 2016, as shown in the time line in Figure 3.1.

The majority (85%) of the complete responses are from female sonographers. The years of experience in sonography of the respondents also varied greatly, with 20-30 years of experience forming the great majority (33%),

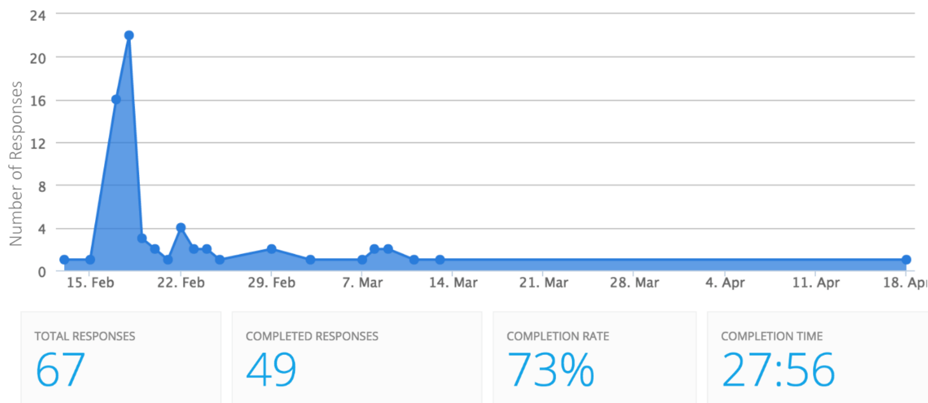


Figure 3.1: Survey Statistics and Number of Responses Over Time

followed by greater than 30 years (23%), 11-20 years (13%), 2-5 years (12%), less than 2 years (10%) and 6 - 10 years (13%), as shown in Figure 3.2.

In terms of current occupation, the great majority of the responses received are from expert sonographers, representing 92% of the survey respondents. The rest are 2 student sonographers, 1 instructor and 1 applications specialist. Based on the results, 77% of the respondents have over 5 years of experience in sonography.

In terms of experience in types of ultrasound exams, the respondents who have an experience in performing general and obstetric/gynaecologic exams form roughly 84% of the respondents, as shown in Figure 3.3. Other ultrasound exams include breast sonography (3 responses), neonatal (2 responses), neuro (1 response), ocular (1 response) and cranial (1 response).

Most of the respondents work for 31 to 40 hours per week distributed as an average of 8 hours a day for 5 days a week, as shown in Figure 3.4.

3.1.3 Interviews

While carrying out the observations and collecting survey responses, structured interviews and informal discussions with two sonographers were being continuously conducted for further clarifications on the observations and survey results and later to assist in the user study design by bringing in a professional's perspective. The first sonographer mainly performs obstetric/gynaecologic (OB/GYN) ultrasound scans and the second sonographer performs general ultrasound scans. Both are expert sonographers with over

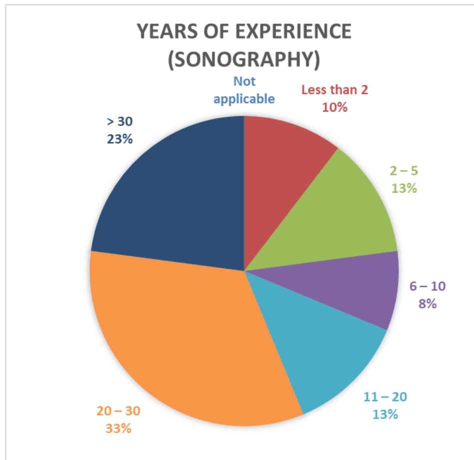


Figure 3.2: Survey Respondents' Years of Experience in Sonography

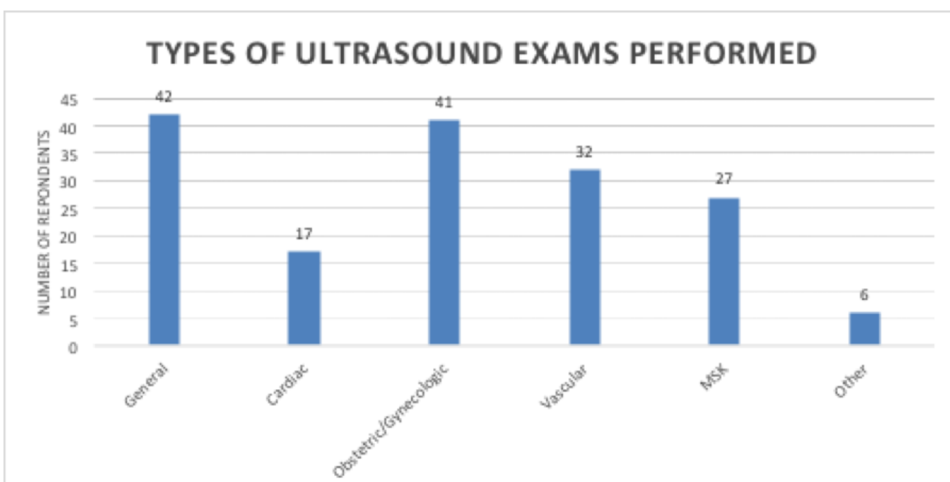


Figure 3.3: Types of Ultrasound Scans Survey Respondents Perform

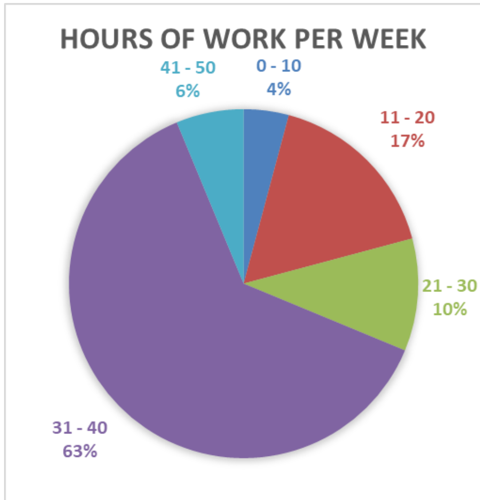


Figure 3.4: Survey Respondents' Number of Hours of Work per Week

30 years of experience in sonography. In addition, we held follow-up discussions with sonographers from the hospitals where the observations took place and with the six recruited sonographers from our user study during the first design iteration.

3.2 Results

3.2.1 Diagnostic Ultrasound Scan Routine

We observed the steps taken by the sonographer during a typical OB/GYN exam through a full-day observation at BC Women's Hospital. Before the exam, the sonographer loads the patient's data, checks the patient's report, and helps the patient lay in the patient's bed. Typically, the sonographer also makes sure the patient can see the ultrasound scan in the secondary monitor.

During the exam, the sonographer typically starts by adjusting TGC, depth and focus levels values after locating the target to be scanned by the probe. The sonographer proceeds to take the required measurements and images are taken during the exam, either to be reported to the physician via PACS or to be printed out for the patient (such as an ultrasound image of the infant on thermal paper). In either cases, the following steps are taken for image capture: (1) freeze image, (2) annotate image, (3) print/capture image, (4) unfreeze. We also observe that the sonographer typically rests her left hand on the freeze button, as it is being frequently used before performing image captures or measurements.

After the exam, the sonographer updates the ultrasound image and data report to PACS. In some hospitals, this is automatically performed through the DICOM communication protocol. Depending on how critical the case of the patient is, the sonographer discusses the results and gives feedback to the patient. If the case is too critical, it is best discussed between the patient and the physician only. The report is then delivered and discussed by the sonographer with the corresponding physician.

3.2.2 Contexts of Attention

As shown in Figure 1.1, a sonographer's attention is distributed across three main visual contexts. The ultrasound image forms the region of central attention, as the goal of a sonographer is to analyze and produce high quality images. Next, the sonographer must concentrate on transducer manipulation, as images must be obtained with the best transducer location. Finally, the sonographers must also use ultrasound machine controls which include various buttons, knobs, sliders, switches, a trackball, and most of the times a touch screen for further menu navigation. In addition, effective communication with the patient is also required, which also accounts for some of the sonographer's attention.

We observed that in obstetric ultrasound scans, it is important for the ultrasound screen to be placed at a location where both the sonographer and the patient are able to see the acquired images. There is typically a second screen placed at a viewing distance from the patient bed, as shown in Figure 3.5.

We noticed that, even the most experienced sonographers, glance repetitively back and forth between the monitor and the ultrasound machine's physical controls. One of the interviewed sonographers reported that the large amount of options sometimes causes unwanted distraction which draws attention away from the ultrasound image. She mentioned a common scenario in obstetrics is repetitively losing the chance to capture the "perfect image" as the fetus rapidly moves while the sonographer is still trying to locate some option on the controls panel.

Based on these observed and discussed challenges with the sonographer, we design relevant questions in the survey, shown in table 3.1, to understand the relevance of these challenges to other sonographers with different experiences.

Figure 3.5: An Ultrasound Room's Layout: A second screen is placed at a viewing distance from the patient for OB scans.

Table 3.1: Survey question 3.1.6 "*Provide your rating of the following. Where 1 = Highly Disagree and 7 = Highly Agree.*"

#	**Statement**
1	Scanning anatomical structures that are in constant motion can be a time-sensitive task that requires an efficient and responsive user interface (such as freeze or print).
2	Ultrasound foot switches can be helpful in repetitive tasks (such as freeze or print).
3	Sometimes I have to go through a lot of steps (through interface menus) to select a particular setting.
4	I switch my attention between the monitor and the ultrasound interface buttons very often and it gets distracting sometimes.
5	I switch my attention between the monitor and the ultrasound interface buttons very often and it makes me lose focus of important image details sometimes.

The results, as illustrated in Figure 3.6, show that survey respondents agreed the most with the first statement regarding the need for an efficient interface to capture ultrasound images of anatomical structures in motion.

Some respondents disagreed that foot switches are helpful for repetitive tasks, but a larger number agreed, which suggests that there is some potential for using foot switches with ultrasound machines, but it is still unclear whether users will welcome this change if it is not necessary.

The third, fourth, and fifth statements, relevant to attention and context switch, show almost an equal amount of agreements and disagreements. This result suggests that there exists a reason influencing the opinion of the respondents, which could be based on the type of ultrasound exams they perform, their experience and their general approach in organizing their work flow.

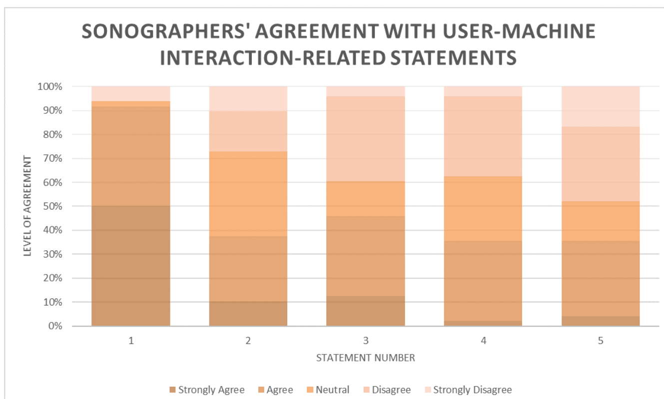


Figure 3.6: Levels of Agreement on the Survey Statements Listed in Table 3.1

3.2.3 Machine Functions and Features

Frequently-used Functions

When asked about their use of ultrasound functions, more than half of the respondents (54%) indicated that they use about 30% to 60% of the ultrasound machine image settings and machine functions out of all the settings and functions they are familiar with, followed by 27% of the respondents who indicated their use of only about 10% to 30%.

We further surveyed the image settings and machine functions used in

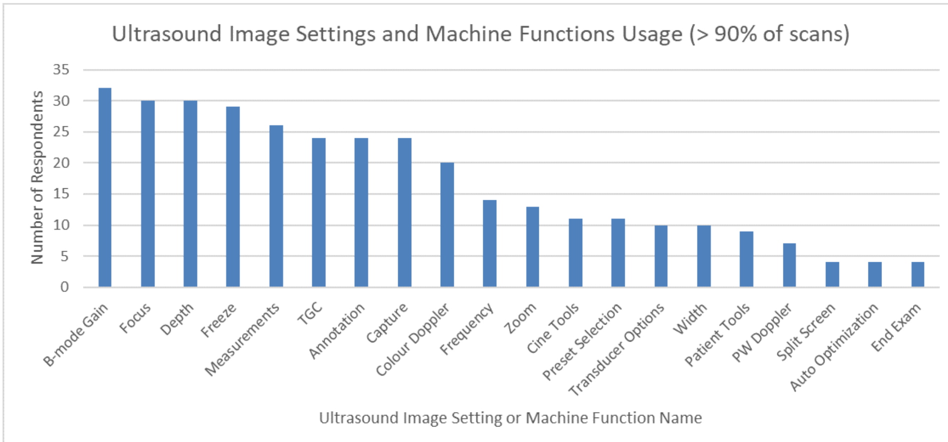


Figure 3.7: Answers to the Survey Question "*Which buttons, functions or features do you use most frequently? at least once per scan in >90% of all scans*"

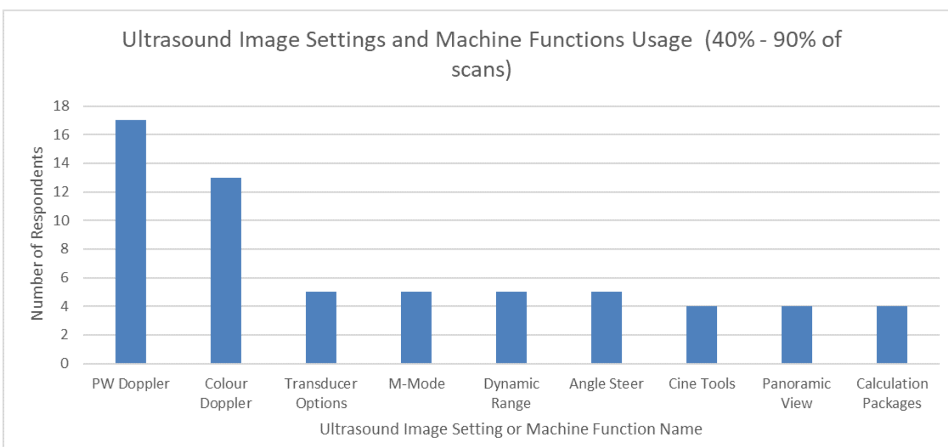


Figure 3.8: Answers to the Survey Question "*Which buttons, functions or features do you use most frequently? at least once per scan in 40% - 90% of all scans*"

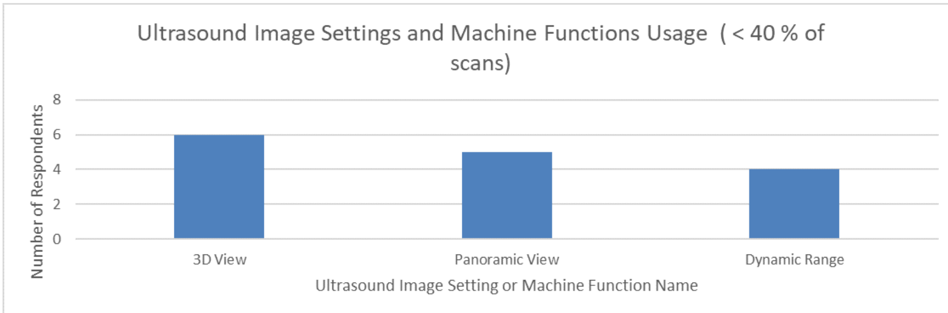


Figure 3.9: Answers to the Survey Question "*Which buttons, functions or features do you use most frequently? at least once per scan in <40% of all scans*"

more than 90% of ultrasound scans. The most common responses are shown in Figure 3.7.

Similarly, we surveyed the ultrasound image settings and machine functions used in some of the ultrasound scans, estimated as between 40% to 90% of the scans performed by sonographers. Figure 3.8 shows the results collected. Figure 3.9 shows the ultrasound image settings and machine features that are rarely used during ultrasound scans.

Adapting to Machine Changes

When sonographers were asked in the survey about the time it takes them to adapt to new changes in an ultrasound machine interface, 48% of the respondents reported that it generally takes them less than one working day to find their way around the machine, followed by 37.5% of them believing that it takes around a week or so. The sonographers' perception of their ability to learn a new interface is a positive indicator to the possibility of introducing slight changes to the ultrasound machine interaction, if its benefits significantly outweigh its drawbacks.

About 10% of the respondents reported other reasons that the learnability of a new interface depends on. For instance, one of the respondents mentioned that it depends on how big the changes are: if the changes are consistent with the general interface of the machine sonographers are used to, then it will not take them long to learn the new additions. Another respondent pointed out that getting used to a new interface depends how often the sonographer should use the newly-added feature. Another respondent mentioned that getting used to a new interface depends on the level of support the sonographer gets.

Evaluation of Machine Features

When asked in the survey about the ultrasound machine hardware and software features that sonographers find helpful, survey respondents mentioned 22 different features.

Touch Screens Although opinions differed regarding touch screens, the majority found distributing the functions between hard and soft keys an advantage, as shown in Table 3.2. One of the respondents mentioned that she prefers soft keys to hard keys in some cases:

"I love touchscreen interfaces. I prefer soft keys to hard keys when they can assist with work flow (i.e. only when contextually necessary)"

On the other hand, another respondent does not prefer this combination as it increases the confusion.

"Machines that have combination of soft keys (touch panel) + hard keys (rotating knobs) + toggles + keys that you press (such as freeze, record, etc.) are the worst. They have too many types of keys, which makes the operation inefficient. They should keep the types of keys to a maximum of two."

Follow-up discussions during the observations at the hospital with sonographers reported that performing menu navigation on the touch screen can get distracting as the sonographer has to look at the menu to perform the settings changes. Physical buttons are always preferred to touch screen buttons, as the sonographer doesn't have to look down at the panel and lose track of where she was looking at the image displayed on the monitor.

Co-located Keys Having all the related controls co-located around the left hand resting area is also reported as very helpful by sonographers, as shown in the ultrasound machine interface in Figure 3.10. We observed that the left hand of the sonographer typically rests around the trackball and the capture and freeze buttons.

Sliding Keyboards On the other hand, when sonographers were surveyed about the ultrasound machine features which they think requires more attention to and improvement, a number of sonographers (N = 7) found the sliding keyboard to be inefficient and difficult to use. The rest of the results can be found in Table 3.3.

Figure 3.10: A Phillips Ultrasound Machine Interface

Table 3.2: The Most Common Efficient Ultrasound Machine Features as Listed by Survey Respondents

Efficient Ultrasound Machine Feature	**# Respondents**
Touch screen functions (automatic annotations, programmable layout, etc.)	11
GE Scan assistant and similar automation protocols	10
Frequently-used buttons are co-located near the hand resting position	8
Keyboard on the same platform as the rest of the buttons	8
Sliding keyboards	4
Patient data loaded automatically	4
Adjustable screen position	3
Pre-programmed annotations	2

Table 3.3: The Most Common Inefficient Ultrasound Machine Features as Listed by Survey Respondents

Inefficient Ultrasound Machine Feature	**# Respondents**
Non-intuitive arrangement of interface panel	13
Sliding keyboards	7
Unadjustable (fixed) monitors	6
Touch panels with multiple layers	4
Hard-to-reach keyboards	3
A lot of steps for a task	3
Sticky trackball / keys	3
Heavy probes, cables and inaccessible ports	3
Hard-to-move machines	2
No leg room under the ultrasound machine	2
Non-intuitive and hard to reach touch screens	5

Scan Automation Another ultrasound machine feature that is found very helpful by sonographers is automation software, such as the GE Scan Assistant, which we dedicate a sub-section to evaluate later in the survey as we found it being used often during the ultrasound scans we observed.

Scan Assistant and Other Scan Automation Software

Out of the 48 respondents, 20 of them have experienced using Scan Assistant in GE machines (or similar software in other machines). "Scan Assistant" is a software created for GE Ultrasound machines that provides "customizable automation at each step of an ultrasound exam for a fast, comfortable, consistent scanning experience, which could reduce injury-causing repetitive actions" [29]. Lowering the amount of repetitive interactions with the ultrasound machine and the amount of time it takes to finish an ultrasound exam are equally leading benefits of such a semi-automated system. Other benefits follow, as shown in Table 3.4, such as helping the sonographer not forget the steps required for a particular exam and focus more on the patient than on the machine interaction. Respondents also reported other reasons for using Scan Assistant, such as preventing the sonographer from incorrectly labelling the ultrasound images.

Although found very helpful, some sonographer survey respondents also reported facing challenges with it, as detailed in Table 3.5.

Table 3.4: Surveyed Benefits of Semi-automated Systems in Ultrasound Machines (e.g. Scan Assistant)

Semi-automated Ultrasound Machine System Benefit	**# Respondents**
Significantly shortens the duration of an ultrasound exam	18
Contributes to lowering the risk of WRMSDs as it lowers the amount of repetitive movements with physical ultrasound machine interface buttons	18
Helps me not forget any steps for a particular ultrasound exam and organizes my work flow	14
Helps me focus more on the patient	11
Other	2

Table 3.5: Surveyed Drawbacks of Semi-automated Systems in Ultrasound Machines (e.g. Scan Assistant)

Semi-automated Ultrasound Machine System Drawback	**# Respondents**
My routine is inconsistent with Scan Assistant	6
Too much automation allows the mind of the sonographer to wander	1
Not optimal for students	1
Slower than manual	1
Can be counter-efficient with lack of training	1

Given this variation in responses, it is important to highlight that such systems work better with some types of ultrasound scans compared to others, which could be the main reason behind the inconsistency of the automated system's routine with the sonographer's routine in some ultrasound scans. When asked about the types of ultrasound scans which the sonographers use Scan Assistant with, carotid ultrasound scans were mentioned by 15 out of 20 sonographers. The rest of the ultrasound scans can be found in Figure 3.11.

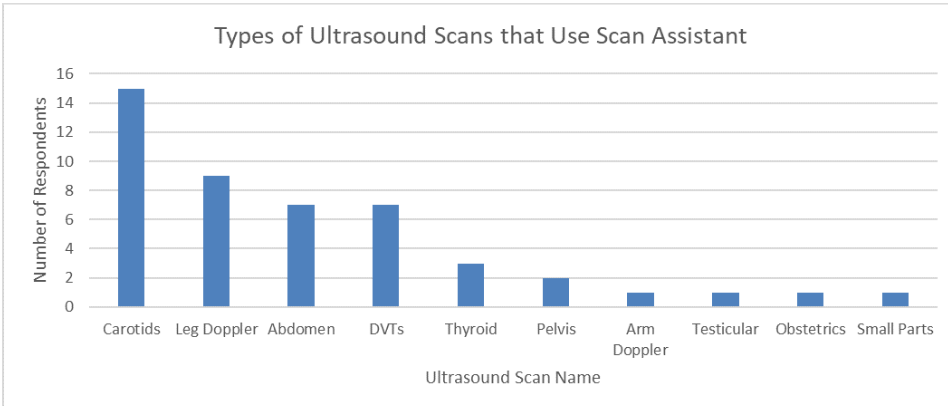


Figure 3.11: Types of Ultrasound Scans that Use Scan Assistant

In addition, one of the interviewed sonographers mentioned that she sometimes has to change the default presets based on some patients' physiology. For instance, scanning obese patients requires a lot of effort to acquire the same level of detail in the image as scanning normal patients, since the sonographer needs to go through changing some particular settings in the default preset to allow the ultrasound signal to penetrate through fat and not cause too much noise.

Evaluation of Voice-enabled Functions

In addition to eye gaze input, voice-enabled systems are another candidate that have the potential to reduce the physical demand of interfaces, if integrated carefully. Old models of Phillips iU22 were the first to integrate voice-enabled functions into the ultrasound machine. We surveyed sonographers to understand their experience with this feature and its advantages and disadvantages within the context of sonography. Out of 48 sonographers, only 12 experienced using voice-enabled ultrasound machines. Respondents found only little advantages associated with this feature, which explains why we do not see it widely used in recent premium to high-end ultrasound machines in the market. Only three sonographers found it helpful, with the advantage being able to reduce repetitive strain injuries. The major disadvantage found (mentioned by 11 sonographers) with this voice-enabled system is the confusion and interference with the sonographer-patient communication as explained by the respondents in details:

"It does not allow you to connect with the patient during the exam. Pa-

tients have less anxiety if you can talk to them and you also need to get more clinical history often."

"I did not like it, I found it distracting for me and the patient and the radiologist if they were present."

The rest of the respondents found it unhelpful due to poor voice recognition or a counter-intuitive interaction:

"It is difficult to learn the entire dictionary required to operate properly."

"Extended exam time due to repeating commands."

On the other hand, one of the time-consuming tasks that the sonographer is required to perform before and after the diagnosis is retrieving the patient's information from and saving to the work list and the PACS system. Based on one of the interviewed sonographers', automating these tasks and introducing improvements to it requiring less interaction time will improve her focus on performing her main job, that is imaging and diagnosis. She suggested using voice commands to automate the retrieval and modification to the patient's data.

3.2.4 Work-related Injuries

There is a great variation in the prevalence of work-related musculoskeletal injuries in the population surveyed, as shown in Figure 3.12, which we assume is related to the variation in age and work experience. Nevertheless, 92% of the respondents reported experiencing stress-injuries they believe is due to their career at least once, which agrees with an earlier study conducted in 2002 regarding the prevalence of musculoskeletal symptoms among sonographers of British Columbia [50].

More than half of the respondents reported that the experienced stress-injuries are rated as either painful (37.5%) or very painful (18.75%), as shown in Figure 3.13, which suggests that the ergonomics of the current ultrasound machine setting require to be seriously considered in terms of redesign and improvement. Looking at what respondents believe the common causes of their work-related injuries are, we find that poor posture is the leading cause, followed by sustained force and pressure, equipment design challenges, repetitive movements, infrequent short breaks and patient obesity. More details are found in Figure 3.14.

Some respondents pointed out more reasons why they believe are the cause of their work-related injuries. These include the following:

- "*Failure of machine manufacturers to provide any sort of support for non-scanning arm/hand*"

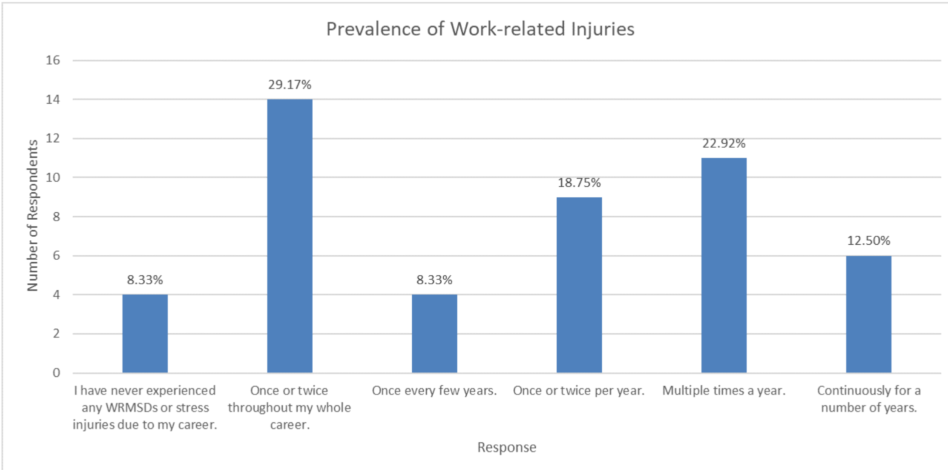


Figure 3.12: Prevalence of Work-related Injuries Among Surveyed Sonographers

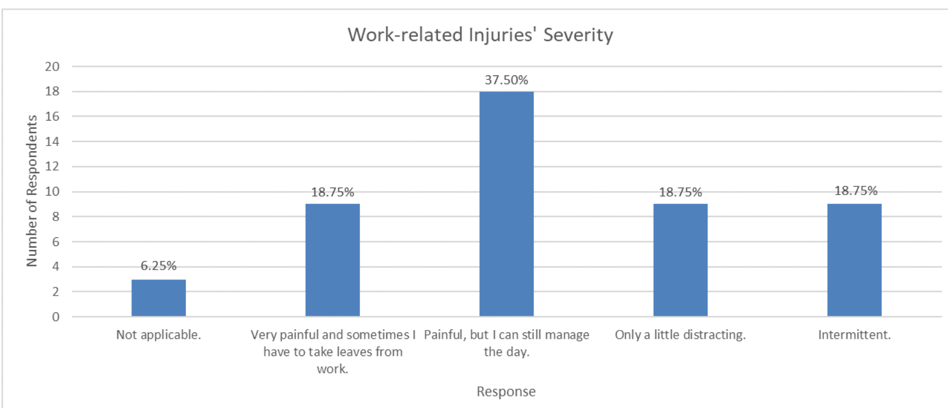


Figure 3.13: Severity of Work-related Injuries Among Surveyed Sonographers

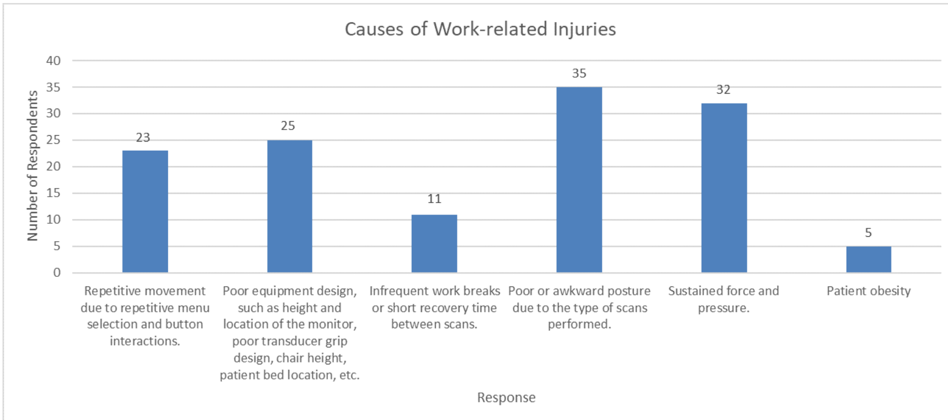


Figure 3.14: Causes of Work-related Injuries Among Surveyed Sonographers

- "*Small fine motor movements*"
- "*High case loads*"
- "*Transducer weight and grip*"
- "*Trackball design*"

One of the respondents explained the difficulty of returning back to work after experiencing a stress-related injury *"Once you are injured at work, even if it is recognized as WCB (rarely), and you are given time off work, you are usually returning to work with pain/muscle weakness and tightness."*

Only one respondent appeared to be confident about his/her inexperience of work-related injuries by reporting: *"I have never had injuries, even when I worked 80 hours a week, part time plus full time job."*

Repetitive Movements

One of the functions we observed in action, panning and zooming, requires the sonographer to perform a number of steps before the image is zoomed. Once the probe is positioned over the required area, the following is performed by the sonographer:

1. Enable the "zoom mode",
2. Using the trackball, move the zoom box to the location of object of interest on the screen,

3. Press a button to toggle the function of the trackball from positioning to resizing the box (or vice versa),
4. Using the trackball again, resize the box,
5. Repeat 2 - 4 to fine-tune the size and position of the box as necessary,
6. Finally, confirm the zoom action.

This is only one form of zooming that is called the High-resolution zoom. Low-resolution zoom is quite straightforward, as it only requires a twist of a knob to zoom in and out and moving the trackball to pan once the image is zoomed. However, sonographers prefer to use the High-resolution zoom as it provides an image with an improved quality and more control over the borders of the zoomed image.

In some scenarios, this repetitiveness in some functions causes some serious interruptions to the ultrasound exam work flow. One of the sonographers recruited for the first-iteration user study, described in Appendix D, reported a difficulty in scanning patients with irregular breathing or those who cannot hold their breath for some time until the sonographer zooms into and acquires an image of the area of interest. Similar to the obstetrics scenario, the structure is also in constant motion due to the patient's breathing. This makes it hard to obtain an image in a certain position while simultaneously performing a number of steps to zoom, freeze and capture the image.

Sonographer's Posture

While observing examinations at BC Women's Hospital, we found that the sonographer was sitting down while performing the first 4 obstetric ultrasound exams, and preferred to stand up for the next 4 gynaecologic ultrasound exams. Therefore, the sonographer's posture while using the ultrasound machine is not always fixed.

In cases where the patients have high body mass index (BMI), we observed that sonographers force themselves awkward and unnatural postures to obtain clear images by simultaneously applying high pressure on the probe and struggling to reach the controls on the ultrasound machine. Sometimes, the sonographer exercises her wrist a little after the exam as the right hand has been exerting pressure on the probe to scan the patient.

3.2.5 Ultrasound-guided Procedures

Physicians who perform ultrasound-guided procedures, are a separate set of users of ultrasound machines. As explained in earlier chapters, their interaction with the ultrasound machine is quite different from sonographers, as they are typically secondary users of the machine and use the ultrasound image to guide them in performing some interventional procedure, and not to optimize and capture images for later analysis. An ultrasound-guided procedure requires the operator's hands to be both occupied by the probe and needle and requires a sterile environment, which prohibits a direct contact with the ultrasound machine's panel to change the image settings. Commonly, most of the settings are preset and their main focus of the radiologist is on the procedure itself (such as needle insertion). However, there might be exceptions.

We included questions regarding a potential hands-free interaction with the ultrasound machine in our survey to get a preliminary idea of its feasibility. Out of the 48 sonographers surveyed, 24 of them have an experience in assisting in ultrasound-guided interventional procedures including biopsies and others, such as porting catheters, thoracentesis and pericardiocentesis.

When asked about the frequency of changing ultrasound image settings during procedures, an equal number of respondents (N = 11) answered "*Yes. All types of procedures frequently require it.*" and"*Yes, but the frequency of this need changes based on the type of interventional procedure being performed.*" Only 6 respondents answered "*No, most of the procedures have image settings pre-set. There might be exceptions though.*" None of the respondents answered "*No, not at all.*", as shown in Table 3.6.

Table 3.6: The Need for Sonographers to Adjust the Ultrasound Machine Parameters During an Interventional Procedure

Need for Adjusting Parameters	**# Respondents**
Yes. All types of procedures frequently require it.	11
Yes, but the frequency of this need changes based on the type of interventional procedure being performed.	11
No, most of the procedures have image settings pre-set. There might be exceptions though.	6
No, not at all.	0

The attending medical staff of the observed breast biopsy included a radiologist and a sonographer. During the observation, we found that the physician performing the biopsy did not interact with the ultrasound machine, except for holding the probe to acquire images at the needed position. The sonographer assisting the physician performed all the ultrasound-related routine starting by capturing images of the biopsy area before and after the procedure and optimizing all the ultrasound image settings. During the procedure, the only ultrasound machine buttons pressed by the sonographer were "freeze" and "capture". Three biopsy samples were taken, therefore three ultrasound images were captured after taking each sample.

Provided assistants are the primary users of the ultrasound machine during interventional procedures, since they perform the required ultrasound image adjustments, we also asked in the survey if there are any difficulties in the communication between the radiologist and the sonographer during an ultrasound-guided procedure. Results are found in Table 3.7. Most of the respondents (N = 11) answered that there is difficulty, but it only depends on the background and training of the assistant. One of the respondents further explained that there is lack of training of assistants recently, which makes this communication quite difficult:

"Intervention is now performed in radiology departments where the support is usually a nurse or a radiology technologist and they have no understanding of the buttons on the unit that need to be adjusted".

One of the respondents explained further that some of the challenging ultrasound image settings to communicate include *"Changing depth, angling the beam, changing a preset"*.

Table 3.7: Difficulty Communicating with an Assistant During Ultrasound-guided Procedures

Need for Adjusting Parameters	**# Respondents**
No, instructions are very straightforward.	5
It depends on the experience and background of the assistant.	11
Yes, but it's tolerable and does not affect the flow of the procedure.	1
Yes, it would be much easier if the radiologist could change the parameters directly.	2

When asked about the potential of integrating a hands-free interaction mode with ultrasound machines for interventional procedures, the majority (N = 11) of the respondents found it helpful, but it does not eliminate the need for an assistant sonographer, as she knows better how to operate the different functions in an ultrasound machine better than a radiologist as she has received complete training on the machine. Results are found in Table 3.8.

Table 3.8: The Preference for Hands-free Control of Ultrasound Machines in Ultrasound-guided Procedures

Need for Adjusting Parameters	# Respondents
Yes, significantly.	7
Yes, but the assistant might still know better in terms of ultrasound machine settings control.	11
No, I do not prefer to interact with the ultrasound machine at all.	2

3.3 Eye Gaze Tracking Integration

When discussing the idea of integrating eye gaze trackers with sonographers, they found that it is worth exploring: as far as diagnostic sonography is concerned, eye gaze input could be very useful, as long as it does not introduce any interaction overhead. Given that the sonographers' attention during a diagnosis session is mostly focused on the ultrasound monitor, providing settings selection with eye gaze tracking (based on where they are looking at on the screen) has the potential to improve their focus on the current task and reduce attentional draw caused by searching for the knobs and buttons on the ultrasound machine's panel.

We find that, although zoom is listed only by 13 sonographers as a function that is used in >90% of the ultrasound scans, it is the first function in the order shown in Figure 3.7 that requires a 2D input, which can be provided as a point of gaze. Another potential function is colour Doppler, as it requires positioning a box around an area of interest to show blood flow. Measurements is another potential function, as it requires placing calipers. However, caliper placement requires high accuracies. Based on earlier research and inherent capabilities of eye gaze tracking, as discussed in Chapter 2, eye gaze input is not a suitable candidate for tasks requiring high accuracies.

Integrating an eye tracker in such an environment can be challenging. Although the lighting conditions in an ultrasound exam room are optimal for eye tracking, calibration is always required for high accuracy eye tracking. Given the typical length of a general ultrasound exam of at least 20 minutes, the routine changes in sonographer positions between sitting down and standing up, the frequent context switch between the monitor and the patient, and in some cases the frequent rotation of the display towards the patient and back towards the sonographer during the exam, there is a risk of calibration deterioration. Another risk for calibration deterioration, as reported in earlier studies [42], is due to a drift effect, which is caused by changes in the characteristics of the eye over time while exerting mental effort due to changes in pupil size or dryness of the eye [44].

3.4 Conclusion

We observed 18 ultrasound scans of various types, surveyed 48 sonographers, and conducted structured interviews and informal discussions mainly with two sonographers throughout the process. The data collected helped us understand in depth the environment of sonography and the sonographers' daily challenges.

Each ultrasound exam is a structured process that involves an understanding of the medical case, continuous communication with the patient, prolonged interaction with the ultrasound machine with the goal of performing scan-specific measurements and obtaining images with an optimal amount of details, and follow-up discussions with the patient and physicians. This process generates three main contexts of attention throughout the scan: the monitor, the patient and the machine controls.

The types of ultrasound machine and image settings and their frequency of use depends on the specific ultrasound scan type. We surveyed other technology that aims at reducing repetitive work-related injuries for sonographers, such as scan automation and voice-enabled machine controls. Semi-automation protocols are widely-used due to their ability to shorten the exam time and reduce potential work-related injuries. Voice commands, on the other hand, are not as efficient as they might seem. This is mainly due to the hindered efficiency of communication between the sonographer and the patient. Nevertheless, we still find a prevalence of work-related injuries (92%) among the surveyed sonographers, which requires attention from hospitals and ultrasound machine manufacturers.

When we investigated the usage of ultrasound machines in ultrasound-guided interventional procedures, we found that the interaction with the machine is minimal compared to sonography. Moreover, it is difficult to replace the role of a sonographer due to her knowledge of the machine operation and the shared cognitive load with the physician performing the ultrasound-guided procedure.

Through this extensive field study, we identified potential advantages to introducing a multi-modal gaze-supported ultrasound machine interfaces, such as a reduced manual interaction with the machine. We also narrowed down the potential ultrasound functions to consider as a starting point for eye gaze tracking integration. We also identified potential risks associated with multi-modal gaze-supported interfaces, such as the need for frequent re-calibrations and the potential interruption of work flow.

Chapter 4

Gaze-supported Interface Design

This chapter presents an overview of the analysis, design and implementation of a gaze-supported zoom interface for the class of ultrasound machines we are investigating. First, we apply the design concepts presented in Chapter 2 related to input device interaction and image browser interaction to zoom and pan functions in ultrasound machines, including High-resolution zoom and Low-resolution zoom. Based on the analysis presented, we present a state-based representation of the zoom and pan functions in ultrasound machines to offer a visual understanding of the interaction. Next, we present an approach for integrating gaze-supported interaction into the presented manual-based state space. We also present our previously-investigated gaze-supported interaction approaches and highlight their limitations. Finally, we present implementation details of the gaze-based features including filtering, simultaneous panning and zooming algorithms and hardware interface design.

4.1 Design Assumptions

As discussed in Chapter 3, currently available ultrasound machines offer High-resolution zoom, Low-resolution zoom and some provide a hybrid function of the two alternatives. However, there is a variation in terms of the image quality produced between the two main zoom functions: Low-resolution zoom post-processes the acquired image to change its magnification, therefore the frame rate is unaffected with zooming and the image quality is not improved. On the other hand, performing High-resolution zoom narrows the area of target acquisition by the ultrasound transducer, thus enhancing the frame rate and providing a zoomed image with higher quality and details. Since High-resolution zoom is not a post-processing function, panning the zoomed image is typically not allowed, as it requires the probe to actively reset the acquired sub-area of the overall visible range. Given these technical differences between the two zoom functions and the diverse set of approaches taken by ultrasound machine manufacturers, we set these design assumptions:

1. **Both Zoom Functions Provide the Same Resolution.** Given that image resolution is not our main concern, we assume all magnification-related functions acquire the magnified area of interest with no improved quality. This means that we are not taking into consideration the adjustment of the probe's lateral and temporal resolution of the image when it is magnified with the High-resolution zoom function.

2. **Rename the Zoom Functions.** Consequent to assumption 1, we refer to the interaction of the Low-resolution zoom in ultrasound machines as "One-step Zoom" (OZ) throughout this thesis and the interaction of High-resolution zoom as "Multi-step Zoom" (MZ).

3. **Enable Pan and Resize.** The sonographer is always able to pan and resize a magnified image, whether it was obtained through OZ or MZ.

4.2 Input Device Interaction Concepts

The layout of ultrasound machines greatly varies based on the target users, available budget and target applications the machine is designed for. Sometimes, machine interfaces differ between manufacturers even if they target the same application and class of users. First, we eliminate a great portion of variety by targeting to analyze the layout interface of only recent premium and high-end machines that are designed for routine diagnostic ultrasound exams performed by sonographers and observed during our field study, described in Chapter 3. This specific class of machines exhibits common characteristics in terms of the input devices on the machine's controls layout and their mapping to the various ultrasound machine functions. Furthermore, we analyze only image magnification-related functions: Low-resolution zoom (or what we refer to as OZ), High-resolution zoom (or what we refer to as MZ), depth and focus. Figure 4.1 shows the layout of the GE Logic E9 ultrasound machine, one of the commonly-used machines in routine diagnostic sonography.

Table 4.1 lists the typical input devices used for the selected ultrasound machine functions. Note that, in some cases, the type of input device is not the same across different machines. Later in this section, we discuss and

Table 4.1: Ultrasound Image Magnification-related Functions and Their Associated Devices

Ultrasound Function	**Input Device**	**Role**
MZ - Multi-step (High-resolution) Zoom	Trackball	Reposition/resize zoom box
	Push button 1	Toggle between trackball functions
	Push button 2	Confirm zoom/reset view
OZ - One-step (Low-resolution) Zoom	Clickable knob	(Turn) Increase/decrease zoom ratio (Click) Reset to original ratio
	Trackball	Pan zoomed area
Depth	Knob or arrow buttons*	Increase/decrease depth value
	Knob or arrow buttons*	Increase/decrease focal depth value
	Push button, knob or arrow buttons*	Increase/decrease number of focal zones

*Type of input device is not consistent across machines for this function

Figure 4.1: The Manual Controls Interface of the GE Logic E9 Ultrasound Machine

suggest the best type of input device based on the presented theory related to input devices.

By referring to Buxton's work on lexical and pragmatic considerations of input structures [10], and our discussions of his work in Chapter 2, we apply the presented theories to the design of magnification-related input devices in this section.

Lexical and Pragmatic Characteristics

Table 4.2 shows four types of dimensionality involved in the sonographer-machine interaction for image magnification-related tasks: 0 (binary), 1D, 2D and 1+1D. The selection of input devices for the binary case can be as simple as a push button. However, the other three types of dimensions require more careful selection by further defining the property sensed and whether it should be mechanical or touch-sensitive.

Property Sensed Given the occupational musculoskeletal disorders already common in sonographers due to their prolonged hours of work with ultrasound machines, using input devices that introduce any type of unnecessary pressure are avoided in the design of ultrasound machine interfaces, if

Table 4.2: Image Magnification-related Functions and Their Tasks Dimensionality

Ultrasound Function	**Task**	**Dimensionality**
MZ - Multi-step (High-resolution) Zoom	Position box	2D
	Resize box	1D (width) + 1D (height)
	Toggle resize and position*	0 (binary)
	Confirm zoom	
	Reset view	
OZ - One-step (Low-resolution) Zoom	Change zoom ratio	1D
	Pan image	1D (width) + 1D (height)
	Reset image	0 (binary)
Focus	Position line	1D (vertical positioning)
	Set number of lines	1D
Depth	Set depth value	1D

*Type of input device is not consistent across machines for this function an alternative can be found. Thus, input devices that sense pressure input are eliminated. In his work, Buxton [10] discusses how to decide between motion-sensing and position-sensing devices by asking the question "*would my input device cause a nulling problem?*" A nulling problem occurs when a position-sensing input device is used for multiple functions in a system, thus changing the value for one function will position the device at a particular place that will interfere with the value of the other function. For instance, if a designer decides to use the same input device for parameter A and parameter B facilitated by a mode switch, the numerical value for both parameters must be the same at all times since the value is directly mapped to the position of the input device, which is not practical. Another source of the nulling problem occurs when the system automatically changes the value of some parameter for auto-optimization purposes or to load some pre-defined settings. In such case, there will always be an inconsistency between the position of the input device that was last set by the user and the value that the system has set. For this reason, motion-sensing devices are the suitable option for ultrasound machine interfaces since the values of image parameters are often reset by the ultrasound machine system or set automatically based on image presets.

Mechanical vs. Touch-sensitive Devices In terms of the last classification property, mechanical-based devices are accessed faster than touch-sensitive devices as they exhibit a more tangible physical interface. Since a sonographer's work requires a lot of context switch, glancing repetitively at the controls to locate touch-sensitive devices (such as touch screens) makes them impractical for image parameters that are used frequently throughout an ultrasound exam and will take up more time and cognitive load by the sonographer.

2-D Tasks For 2-dimensional tasks, we are presented with the choice between a trackball and a mouse. A number of studies prove the superiority of mice over trackballs in terms of time and accuracy for achieving pointing and dragging tasks, such as the study conducted by MacKenzie *et al.* [43]. However, a follow-up study by Kabbash *et al.* [35] that adopted the same tasks and input devices for their user study took into consideration the differences in the performance between the dominant/preferred and the non-dominant/non-preferred hand. In their study, they tested the hypothesis *"Preferred and non-preferred hands yield the same speed, accuracy, and bandwidth using the mouse, trackball and stylus in pointing and dragging* *tasks"*, which they rejected as the use of the trackball by the non-preferred hand yielded higher accuracies than the preferred hand. Additionally, they argue, "*The ease of acquiring a fixed-position device (such as a trackball, touch pad, or joystick) may more than compensate for slower task performance once acquired.*" In the case of the multitude of functions a sonographer performs with ultrasound machines which requires a detailed layout of physical input, using a non-stationary input device, such as a mouse, will require them to lift off their arms repeatedly to reach it and the rest of the physical inputs, which consumes extra time, effort and cognitive load that could be avoided by using a stationary input device. Given that it has been empirically proven that a trackball outperforms a mouse in terms of accuracy when used by the non-dominant hand, trackballs are an integral part of all current ultrasound machine interfaces for 2-dimensional tasks. Another input device which has similar capabilities to a trackball and has not been discussed by Buxton [10], is the touch pad with multi-touch features, enabling seamless zooming and panning. However, we have not observed any ultrasound machine interfaces with an integrated touch pad. Therefore, it is outside the scope of our discussion.

1+1D Tasks Using a 1D device with a mode switch is sufficient to perform a 1+1D task. For instance, a clickable continuous rotary knob can be used to set the width and height of the area of interest in an image. Another approach would be to use a 2D device with a threshold. The latter approach is taken when the number of input devices on a layout needs to be minimized, where the same 2D device can be used to perform a multitude of functions including 1D, 1+1D, and 2D tasks. In ultrasound machines, a trackball is often performing 1+1D involved in magnification-related functions.

The Three-state Model

Based on Buxton's three-state model [11], the types of transactions performed by a sonographer for each zoom function fall under the point/select tasks. We generated the three-state model for both Multi-step (High-resolution) zoom and One-step (Low-resolution) zoom shown in Figures 4.2 and 4.3. Depth and focus are trivial cases with one state only for each input device involved, which we will briefly discuss.

Figure 4.2 shows the three-state diagram for the input devices involved in performing OZ in ultrasound machines including the clickable knob and the trackball. Figure 4.3 shows a more complicated case for performing MZ for the input devices involved in performing the function, that is the trackball and its supporting push buttons.

One-step (Low-resolution) Zoom Figure 4.2 describes the interaction with the OZ function that is composed of two independent user transactions. Referring to Table 4.2, OZ is composed of three tasks: one is a 1-dimensional task that handles the magnification ratio and another is a binary task that resets the magnification. The third task is independent of the other two tasks as it handles panning the magnified image. A suitable input device for the 1-dimensional and binary tasks is a continuous clickable rotary knob. The system is always in state 1 tracking the knob's position, unless it is pressed, which initiates a state 1-2-1 transaction. The third task involved in the OZ function is a 1+1D task, which, as explained earlier, can be realized with a thresholded 2D input device such as a trackball. Referring to Buxton's work [11], if no other input devices assist the transaction performed by the trackball, it exhibits only one state: the tracking state. Therefore, the three-state diagram is at its simplest case.

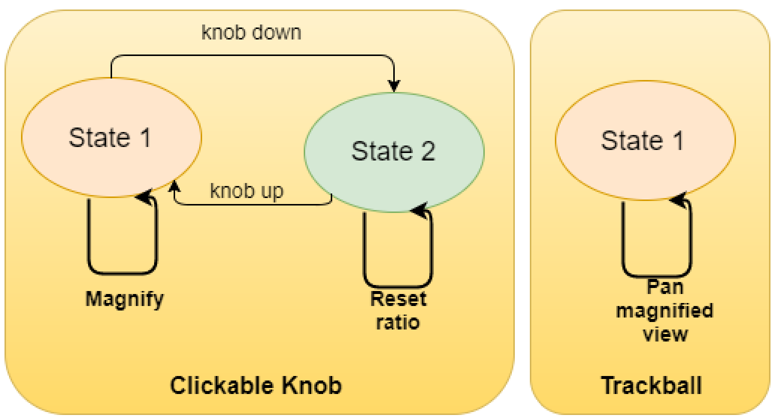


Figure 4.2: Three-state Diagram for One-step (Low-resolution) Zoom in Ultrasound Machines

Multi-step (High-resolution) Zoom 2D, 1+1D and three binary tasks compose the interaction with the MZ function. Note that two of the binary tasks (confirm zoom and reset zoom) occur in different image modes; therefore the same binary input device (a push button) can be used for both. As discussed earlier, a trackball is the selected choice for performing 2D tasks with ultrasound machines. The selected input device(s) for the 1+1D task would determine the number of binary tasks (and therefore the number of input devices) there are to support the rest of the function. The following are the possible options:

1. Two separate 1-dimensional input devices (width and height control of the zoom box), one 2D device (reposition the zoom box) and two binary input tasks (confirm and reset zoom),

2. One 1-dimensional input supported by a mode switch (toggle between width and height control), one 2D device (reposition) and three binary input tasks (confirm zoom, reset zoom and mode switch for the 1+1D task), or

3. One 2D device with a mode switch (toggle between resize and reposition the zoom box) and three binary input tasks (confirm zoom, reset zoom, toggle between resizing and repositioning).

The diagram in Figure 4.3 shows the full interaction of Multi-step (High-resolution) zoom. Note that there are two main modes/states in this interaction: Pre-zoom and Zoom. The Pre-zoom mode contains the set of tasks a sonographer performs to set the size and position of the area of interest, while in the Zoom mode, the only possible interaction is to reset the image to its original view and to pan the image with the trackball. Given this mode switch, a single push button can be used to perform the tasks confirming the zoom and resetting the image. In addition, within the Pre-zoom mode there are two sub-modes: reposition and resize. In the diagram, we denote Pre-zoom's reposition with (a), Pre-zoom's resize with (b) and Zoom with (c). Thus, states 1 and 2 can be in any of the three (sub)modes based on the user's interaction.

Performing a composite press and move with the trackball lowers its performance, as investigated by some studies such as MacKenzie *et al.*'s [43], which attributed the decrease in performance to the interaction between the small muscle groups of the fingers while moving the trackball and simultaneously holding a button. Therefore, similar to One-step (Low-resolution) zoom, state 2 is always a 1-2-1 transaction.

When the trackball is in state 1(a), two possible options can be performed with the two different mode switch input push buttons: either to toggle to resize the area of interest or to confirm the zoom. Similarly, if the sonographer decides to toggle to resize, the trackball's state changes to state 1(b) where the there are two possible options with the same mode switch input push buttons: either to toggle to reposition the area of interest, which

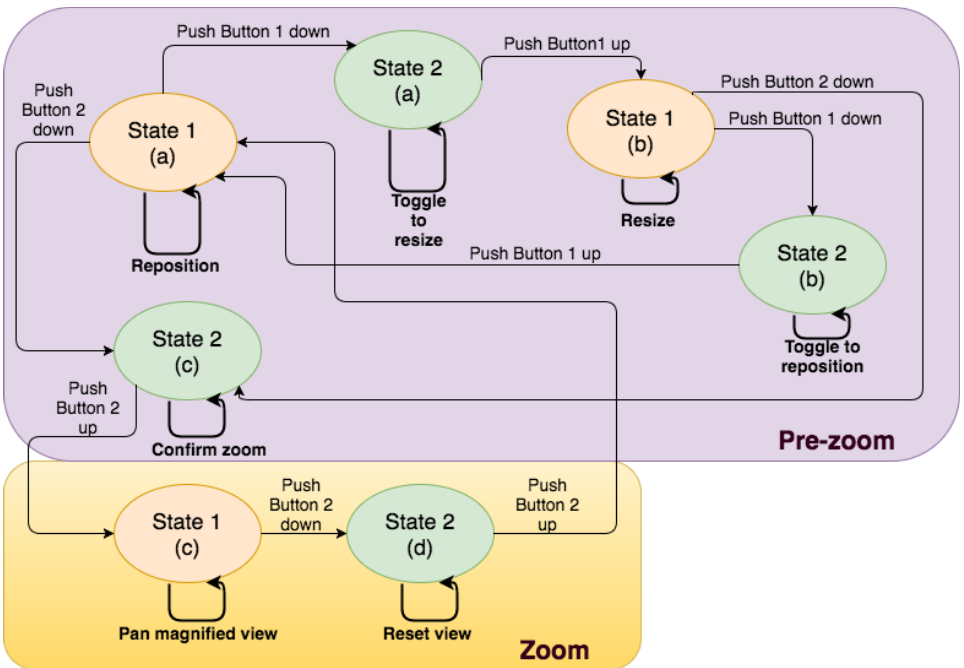


Figure 4.3: Three-state Diagram for High-resolution Zoom in Ultrasound Machines

takes the trackball back to state 1(a), or to confirm the zoom. If the sonographer decides to confirm the zoom while the trackball is in either of states (a) (reposition) or (b) (resize), the trackball moves to state 1(c), where it is able to pan the magnified view, similar to its function in OZ.

Focus and Depth Depth is another function in ultrasound machines that magnifies the ultrasound image by adjusting the transducer's image acquisition to display a deeper view of the target. Although we do not extensively analyze it, we classify depth as a magnification-related function. Similarly, focus is not particularly a magnification function. However, we include it in our discussion on image magnification-related functions in ultrasound machines, as it is one of the direct image features that are changed along with magnification. In ultrasound images, the focus is the horizontal line where the set of ultrasound beams produced by the probe are all focused at, therefore producing the highest lateral resolution at that depth. The majority of the recent machines support multiple focal points, so an ultrasound operator can set a number of focal depths where the lateral resolution of the image is highest. However, increasing the number of focal points comes at the cost of decreasing the temporal resolution.

Often, when a sonographer changes the image magnification, especially in MZ, the focus is reset to some position within the magnified area. Most machines place the focus in the middle of the magnified view. By referring to Table 4.2, focus is composed of two tasks: reposition the focal point(s) and change the number of focal points. Both tasks are 1-dimensional, so the optimal input devices for these tasks would be either a continuous rotary knob or a continuous treadmill thumb-wheel. In both cases, the three-state representation is a trivial 1-state diagram where the input device is constantly tracking the position of the input device. The implementation in ultrasound machines in terms of the choice of input devices for focus differs from one machine to another. However, most implementations dedicate a knob input solely for changing the position of the focal point(s), as it is frequently accessed during a routine diagnostic ultrasound exam and place the option for changing the number of lines within the touch screen display.

Similarly, depth is a trivial 1-state case diagram where the system is constantly tracking the position of the rotary knob or thumb-wheel dedicated for setting the depth of the image. Depth is one of the most frequently-accessed ultrasound functions; ultrasound machines typically have a separate dedicated input device for depth setting, as shown in Figure 4.1

4.3 Image Browser Interaction Concepts

In this section, we explain the approaches taken in the design of the magnification-related functions in commercially available ultrasound machines based on Plaisant *et al.*'s [47] image browser classification theory and compare and analyze the choices taken based on the tasks and the class of users interacting with different image browser alternatives.

In routine diagnostic sonography, a sonographer's main task is to locate and acquire an anatomical target and then to optimize the acquired image to clearly show a particular area of interest. Often, a sonographer is also required to perform secondary post-acquisition tasks such as measurements and diagnostics and report all the acquired, generated and measured data to a physician. Based on Plaisant *et al.*'s [47] classification, a sonographer mainly performs diagnostic tasks on the image in an image browser. Sometimes, a sonographer's task might also be classified as open-ended exploration as the sonographer is still in the localization phase of the anatomical target to be acquired in the image (for example, as the sonographer is locating the liver in an abdominal ultrasound exam).

Once a target's location is acquired with the probe, a sonographer might change the depth of the image to bring the target to the central attention or perform a local zoom operation to be able to accurately measure an area of interest. In either case, the sonographer's task is a diagnostic task as she is investigating an area of interest within a magnified image. For diagnostic tasks, speed of panning and zooming is an essential requirement for the image browser's interface since the sonographer's short-term memory is actively being used to compare patterns and investigate parts of the image. Additionally, the interface should include a mechanism to provide the sonographer with a context within a zoomed image, as complete coverage is crucial to show the target in relation to the surrounding anatomy.

Presentation Aspects Figure 4.4 shows the complete taxonomy for browser presentation aspects as presented by Plaisant *et al.* [47], with additional check marks next to the optimal presentation aspects for ultrasound machine image browsers. As mentioned earlier, an effective browser interface has to provide the sonographer with contextual awareness. Therefore, a hybrid of single and multiple views is a rational design approach for the presentation of the image browser in ultrasound machines, where the multiple views are presented as an overview-detail pair where the detail acts as a zoom-and-replace view and the overview provides the sonographer with the dimensions and the location of the detailed view. Given that the sonogra-

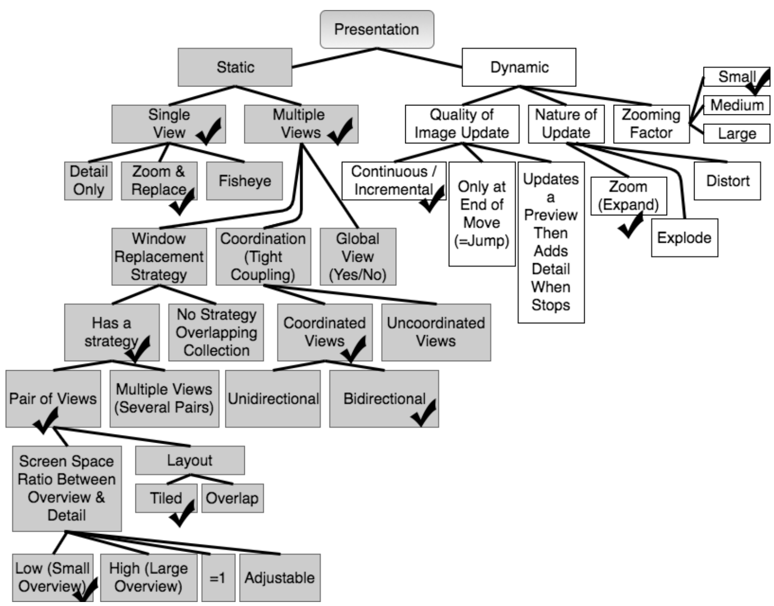


Figure 4.4: Browser Taxonomy Presentation Aspects [47]: A check mark is added next to the design choices for ultrasound machine image browsers.

pher performs accurate measurements on the detailed view after zooming into an acquired image, the aspect ratio between the overview and detail must be low, with a larger detail view since it is the central attention of the sonographer. Consequently, any changes within the detailed view are mirrored to the overview, and vice versa, making the views tightly coupled with bidirectional coordination, which reduces the cognitive load of the sonographer by not having to manage the views in addition to managing her primary ultrasound-imaging task. Similarly, continuous/incremental update of the image (in this context, by "update" we refer to updating the detailed view as the sonographer zooms in) allows the sonographer to concentrate on her main ultrasound image acquisition task and not on the browser navigation tool. In the presented taxonomy, the authors [47] classify the nature of the image update to zoom, explode and distort. Due to the nature of images and the transducer capability of ultrasound machines, the only applicable option for updating an ultrasound image to reveal higher definition detail is magnification-related functions (zoom and depth). Explode refers to revealing an internal structure of the zoomed part, which is applicable in areas such as dense maps or network diagrams. Distort refers to a fisheye-like presentation, such as the work presented in [56], which is not suitable for ultrasound images as it will interfere with an accurate presentation of the structure and with accurate measurements of the target. The zooming factor is the ratio between the presented image in the overview and the detailed views in terms of the level of magnification. Plaisant *et al.* [47] suggest empirical testing to set these values as they differ from one application to another. However, based on their experience, the ratio between the two views should not exceed 20:1, otherwise intermediate views should be called for. In ultrasound imaging, high zoom levels are not required as in the case of dense images, as the magnification does not reveal new data, but only centres the target of interest in the view and allows for higher accuracies of operations to follow. For instance, GE's Voluson E8 does not allow a magnification ratio beyond 3.4:1.

Operation Aspects Figure 4.5 shows the operation aspects taxonomy based on Plaisant *et al.*'s [47] work. We modified it by adding extra options to zoom and pan as implemented in the ultrasound machines' image browser interface. Similar to the presentation taxonomy, we highlighted the operation approaches followed in ultrasound machines. However, due to the diversity of machines in the market, there is no unified approach to implementing the operation aspects of ultrasound image browsers. As mentioned

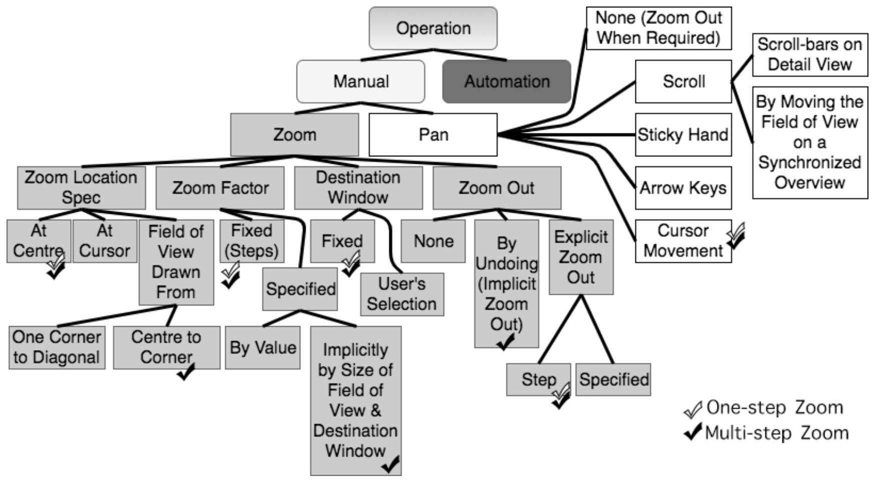


Figure 4.5: Browser Taxonomy Operation Aspects [47]: The design choices for the different alternatives of ultrasound machine image browser magnification-related functions are highlighted with different check marks.

earlier in this chapter, most machines offer a Low-resolution quick-access zoom function (OZ) and another High-resolution zoom (MZ) function that provides the machine operator with higher control over the location and dimension of the magnified image, which requires the usage of more than one type of input device to achieve the magnification task. Some newer machines offer a hybrid of both zoom options, which some machines name it the PanZoom option: a Low-resolution magnification with extra control options over the area of interest's dimensions. Secondly, machines differ especially in the design of the operation of the High-resolution zoom (MZ). Through our market research and field observations, we were able to find the design trends in the image browser zoom design for premium and high-end ultrasound machines used in routine diagnostic ultrasound exams. Finally, sonographers often interchangeably use depth and zoom as these functions serve the same task with varying levels of control and image quality to magnify an area of interest.

In Figure 4.5, we separately identified the operation aspects followed by High-resolution zoom (MZ) and Low-resolution zoom (OZ). Additionally, we omitted the automation recommendations of the operational aspects of image browsers, since we are concerned with the design of the manual magnification function itself, and not in secondary image browser tasks discussed by Plaisant *et al.* [47] such as saving points, navigation, window management and search.

State-based Analysis

To understand the mapping between ultrasound machine manual input devices and their functionalities for performing the zoom and pan actions, we generate a state-based analysis. This analysis also provides a visualization of the interaction that helps later in identifying the potential areas of eye gaze input integration.

We focus on zoom functions that implement three states: Full-scale image, Pre-zoom and Zoom.

1. **Full-scale** is the state where the whole image is displayed to the sonographer and panning is not possible.

2. **Pre-zoom** is the state where the user is actively changing the dimensions and position of the zoom box, or area of interest to be zoomed into.

3. **Zoom** is the state where only a magnified portion of the whole image is visible.

Figure 4.6 shows the typical interface layout used to operate magnification-related functions. In addition, Figure 4.7 is a state diagram for the zoom functions discussed. An illustration of the interaction is shown in Figure 4.8.

The Zoom and Pre-zoom states each have two sub-states. Based on the sub-state, the functions of some of the input devices shown in Figure 4.6 change. In sub-state (a), in both Zoom and Pre-zoom states, is associated with movement:

- **Zoom (a)** the trackball pans the image
- **Pre-zoom (a)** the trackball repositions the zoom box.

In sub-state (b), the function of the trackball is to resize:

- **Zoom (b)** the trackball resizes the zoomed area
- **Pre-zoom (b)** the trackball resizes the zoom box.

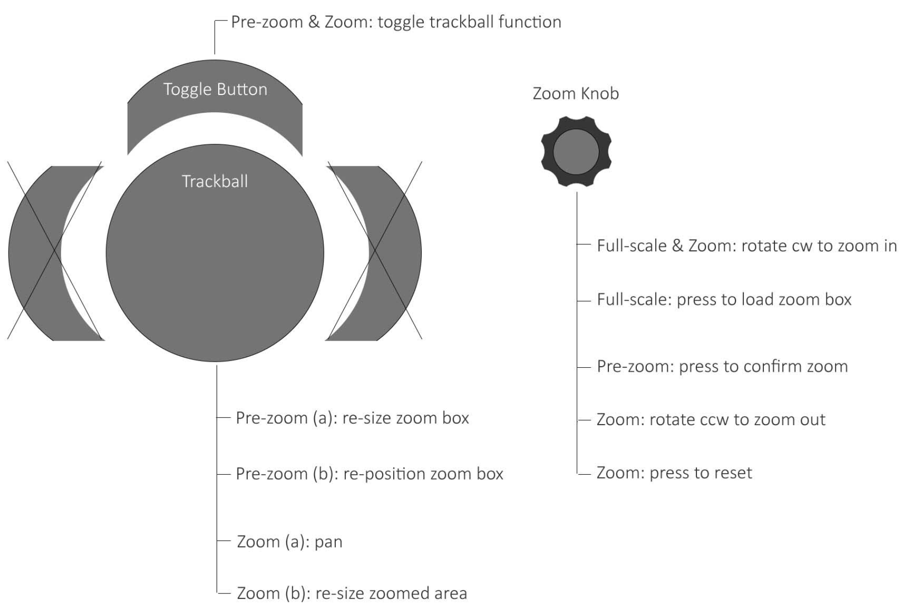


Figure 4.6: The Input Layout Design of the Traditional (Manual-based) Ultrasound Machine Interface with the Mapped Functions per State for Zoom Functions.

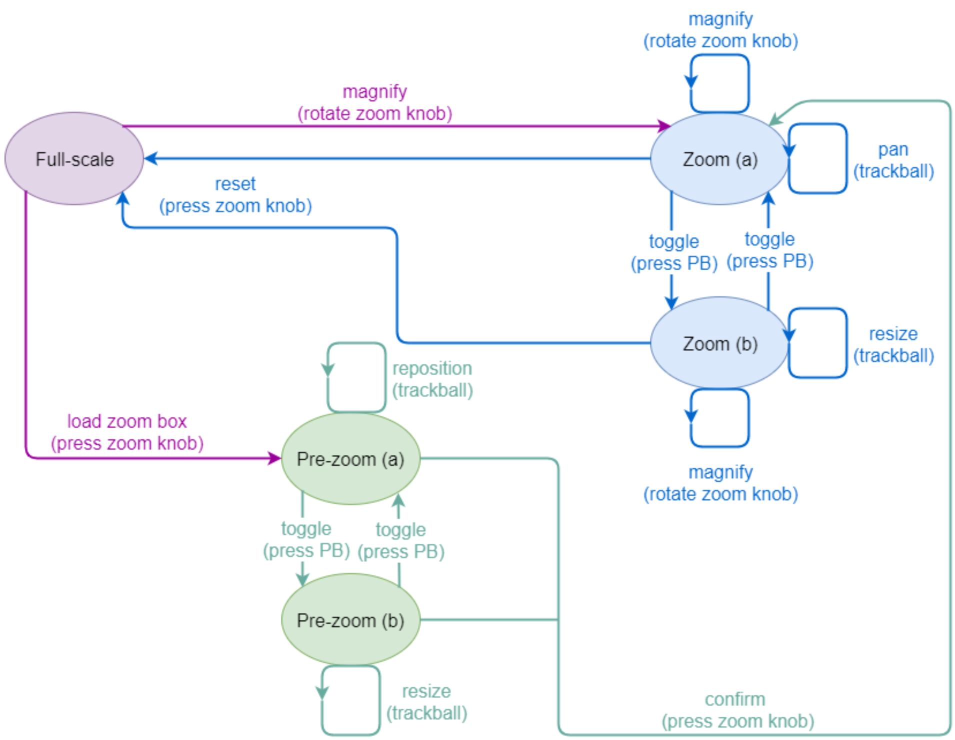


Figure 4.7: A State Diagram of Zoom Functions in Ultrasound Machines: MZ zoom includes all three states. OZ includes only the Full-scale and Zoom states.

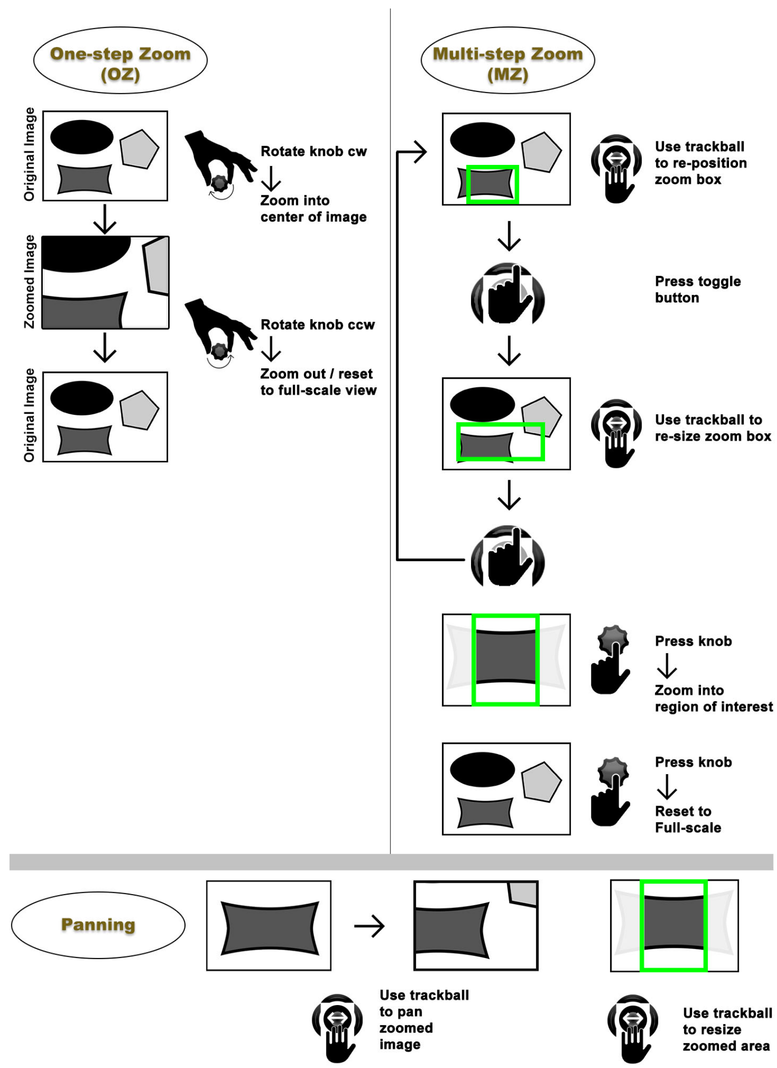


Figure 4.8: An Illustrated Interaction of One-step Zoom, Multi-step Zoom and Panning of Traditional (Manual-based) Ultrasound Machine.

Switching between the sub-states, (a) to (b) or (b) to (a), is always done with the same toggle button.
In the ultrasound machine layout design, the zoom knob's button is the main input for transitioning between the states in a periodical fashion: in the Full-scale state, pressing the button transfers to the Pre-zoom state, in the Pre-zoom state, pressing the button transfers to the Zoom state, in the Zoom state, it transfers back to the Full-scale state.
The zoom knob rotation is only enabled in the full-scale and zoom states: it zooms in or out with a constant ratio and uniform image border size.

OZ includes only the Full-scale and the Zoom states and sub-states. MZ includes all three states.

4.4 Integrating Eye Gaze Input

By mapping eye gaze to Buxton's [10] tableau of input devices, eye gaze serves as a 2-dimensional input that conveys position. Any other actions to be performed by the user related to selection must be conveyed either through another channel (such as manual input, speech, gesture, etc.) or through other eye behaviour, such as dwell time, in case a multi-modal option is not feasible. Consequently, using the three-state model [11], independent eye gaze input supports states 0 and 1, where it can be either OOR (eye gaze is undetectable by the tracker), or being actively tracked for position. A second input modality, such as manual buttons, must be integrated for eye trackers to support state 2.

In his work, Jacob [32] classifies the interaction techniques performed with eye tracking-supported interfaces. One or more interaction techniques can be present in an interface based on the task requirements of the system. The first technique discussed is "*object selection*". As the name suggests, object selection is the action of intentionally looking at a particular object in the user interface and performing selection. Similar to object selection is another interaction technique called *"continuous attribute display"*, which is the retrieval of information of the particular object on the screen upon user selection. *"Moving an object"* is another interaction technique, which can be achieved in one of two approaches, where in both the eye is used to point at the particular object of interest to be moved. In the first approach, the manual input signal performs both the confirmation of selection and moves the object, while in the second approach, the manual input is used only to confirm the initial and final positions of the object and the eye is used to drag the object (the object is latched to eye movement). The next interaction technique is *"Eye-controlled scrolling text"*, which can be generalized to scrolling of any type of content within graphical borders. The idea behind it is that as the eye reaches the end (edge) of text (content), the interface naturally scrolls to reveal the rest of it. *"Menu Commands"* is another interaction technique discussed, which is quite similar to *"object selection"*, except that the objects being selected are graphical user interface menu items. The final interaction technique discussed is *"Listener Window"*, which automatically sets the listener/active window based on the user's location of gaze in window-based graphical user interfaces.

In both our design alternatives, we follow the recommendations made by [68] and followed by numerous gaze-supported interaction work, such as [42], to not overload the visual channel with motor control by enabling eye gaze input only when combined with a manual input to initiate a motor control action.

4.4.1 One-step Zoom

One-step zoom could leverage eye gaze input implicitly by using the point of gaze to define the area of interest the user aims to zoom at. Traditionally, OZ magnifies the centre of the image. If the user were interested in an area that is located towards the side of the image, manually zooming into the centre would require the user to eventually pan the zoomed area to reach the area of interest located at the side. With eye tracking, the system is already aware of the area of interest the user aims to zoom at, therefore it would require less or no follow-up panning to reach the area of interest. From an abstract perspective, this interaction falls under *"object selection"* [32], since the sonographer implicitly selects a particular feature in the image to zoom into/out of it.

4.4.2 Multi-step Zoom

By referring to Jacob's [32] work, this interaction corresponds to *"moving an object"*, as the eye gaze input is used to set the position of an object on screen. In this context, the object is the zoom box that defines the area of magnification. The work presented in [32] suggests two techniques: the first is to use the eye gaze only to select the object to be moved then use a manual input device to perform dragging, and the second is to use the eye gaze to select the object and latch the position of the object to the eye gaze as long as another muscle group is engaged in the interaction (a button is depressed, for example). In the case of our application area, there is only one object to be selected, which is the zoom box, so the first implementation approach would not apply. Furthermore, in [67], they observe that users preferred the second approach more as once they pick up the object, they directly look at the desired destination, and thus latching the position of the box to the eye gaze would make the interaction much faster. They also note that users performed better when the destination formed a clear recognizable feature, and not just white empty space, as it is easier for the eye to look at and fixate on particular features. However, one unavoidable risk that the interaction could run into due to a moving object on the screen latched to the eye gaze is the positive feedback loop, as explained later in Jacob's [32] work on eye tracking interfaces, which might occur in case the initial calibration was not accurate enough, which will cause the user's eyes to be drawn to the object and displacing it further.

Although we cannot employ the same interaction techniques explored by Zhai *et al.* [68], as there are no distinct targets in the interface known to the system as "hotspots" and the whole image is treated as a target, we could still adopt some concepts, especially from his conservative design approach, which warped the eye gaze to the targets and fine movements are further performed by manual input. In our design, the zoom box is always warped to the user's fixation and fine details, such as the box dimensions are determined by a manual input.

This design follows the first refinement step in the two-step refinement zoom approach described in [42]: the region of interest is only defined within a confidence interval (zoom box) as the user presses a key and looks at a specific area. Contrary to [42], our design does not capture a zoomed area within the zoom box before performing the zoom action, as this will distort the underlying ultrasound image. The full two-step refinement approach is more suited for fine targets on screen, such as for text reading or web browsing.

4.4.3 Gaze-based Panning

Following the eye-gaze interaction approach introduced by Jacob [32] named "eye-controlled scrolling text", eye gaze input can also be used in our context to scroll the zoomed image to reveal more content, just as it is presented in Jacob's [32] work (and many later studies, such as [34]) to reveal more text once the point of gaze approaches the edges of the image.

4.5 Proposed Design

Through exploring a number of design alternatives, as detailed later in section 4.6, we arrive at our final interaction design, as shown in the layout diagram in Figure 4.9 and in the state diagram in Figure 4.10. An illustration of how this interaction works is further graphically illustrated in Figure 4.11.

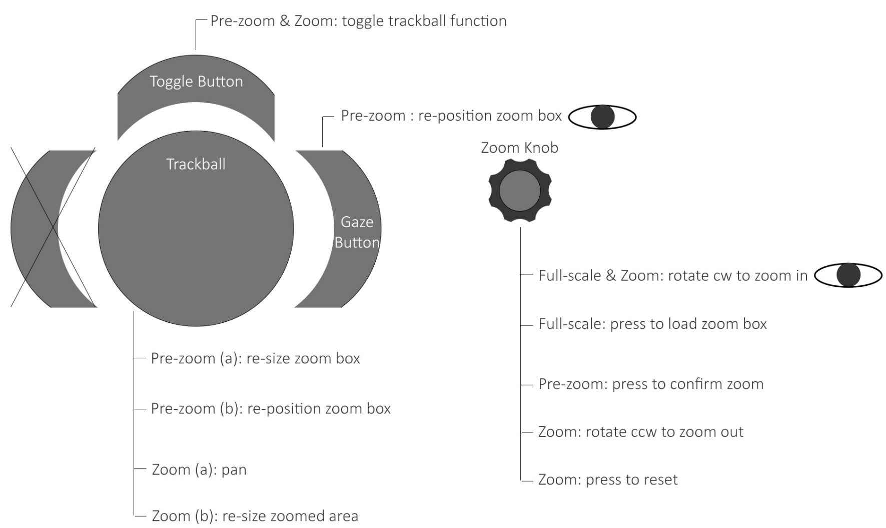


Figure 4.9: The Interface Layout of the Proposed Design Alternative: the active input elements are the trackball, the toggle button, the gaze button, and the clickable zoom knob.

This design has the same base as the interaction design of the typical ultrasound machine zoom functions explained earlier in this chapter. The differences are the following added gaze features.

1. In the Zoom state, rotating the zoom knob (zooming in) always takes the point of gaze as an input and zooms into that area.

2. In the Pre-zoom state, holding a button moves the zoom box based on eye gaze input.

The first feature targets the One-step Zoom function by using the user's eye gaze implicitly as the user is already looking at the point of interest

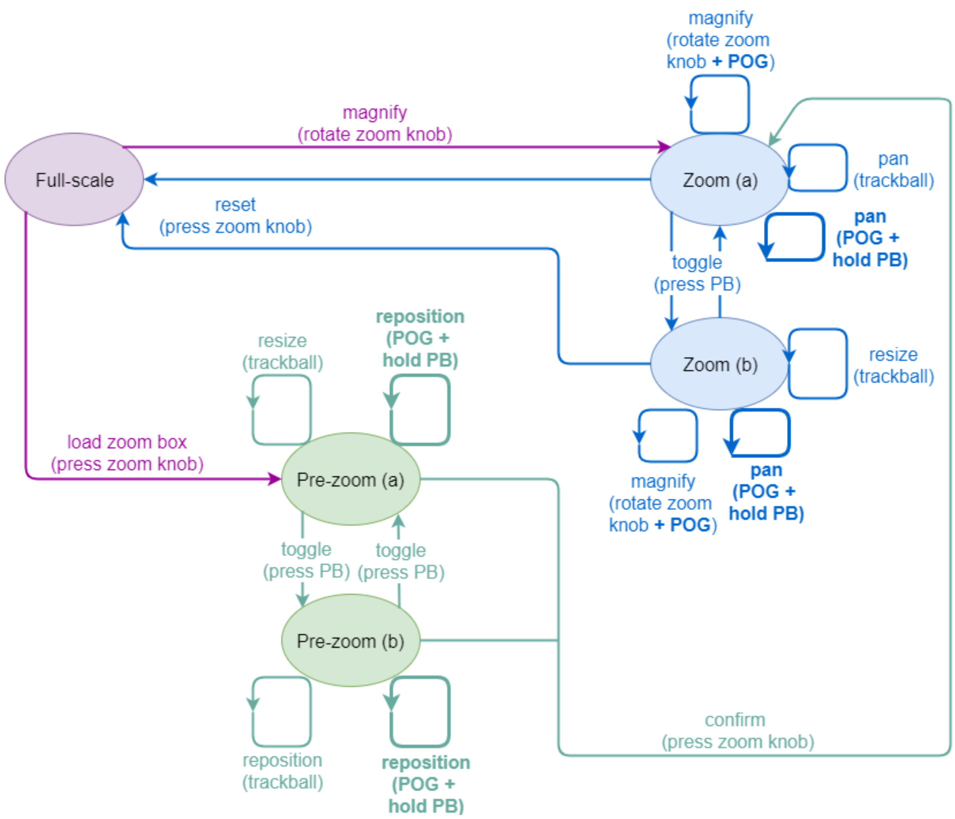


Figure 4.10: The Interaction State Diagram of the Proposed Design Alternative: the same as the state diagram in Figure 4.7, with four added gaze-supported interactions taking the Point of Gaze (POG) as an input. In the Zoom states, eye gaze input is implicitly integrated. In the Pre-zoom states, eye gaze input is explicitly used to move the zoom box.

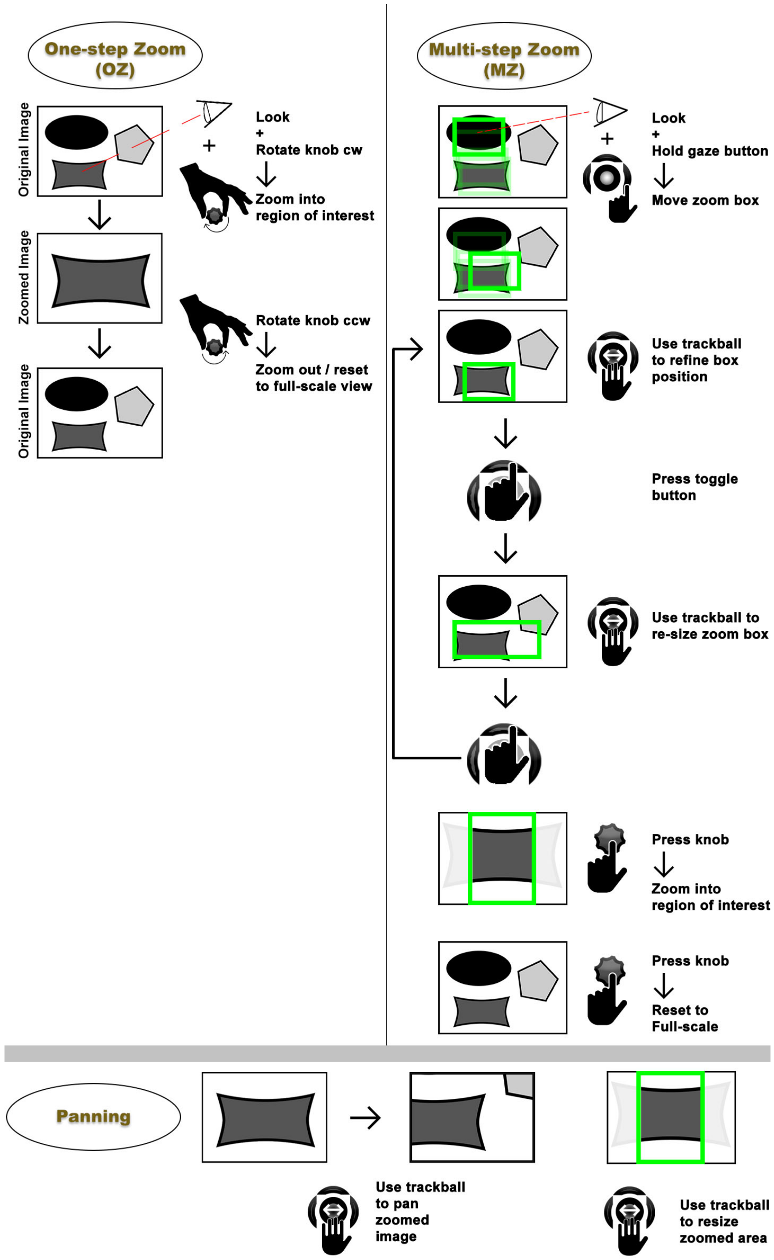


Figure 4.11: An Illustrated Interaction of One-step Zoom, Multi-step Zoom and Panning of the Proposed Design.

while zooming. In the manual-based interaction alternative, One-step Zoom requires iterative panning as zooming in always uses the centre of the visible image as the point of interest. In this gaze-supported alternative, the need for panning is omitted, as the system automatically sets the point of gaze within the visible image as the point of interest. This is designed to reduce the manual interaction and speed up the zooming task.

The second feature targets the Multi-step Zoom function by latching the movement of the box to the movement of the user's eye gaze. Although this is an explicit gaze-supported interaction, it is only activated upon pressing and holding a button, which is designed to eliminate the Midas touch problem discussed earlier in this thesis. Similarly, this is designed to reduce the manual interaction by reducing the use of the trackball to move the zoom box and consequently speed up the interaction.

Unlike earlier work in the field [25] [2] [46], we did not integrate gaze-supported panning due to negative feedback and observations during pilot studies. Similar challenges have been reported in earlier work [54] [30], where participants found the gaze-based panning feature interfering with visually tracking the object.

4.6 Earlier Investigated Design Alternatives

In this section, we present the earlier design alternatives we investigated for One-step Zoom, Multi-step Zoom and gaze-supported panning and explain the reasons they proved to be inefficient either through tests performed with end users or through an analysis of interaction.

4.6.1 Alternative 1

This design alternative for OZ, MZ and panning in both zoom techniques was designed and tested with end users during the first iteration of this user-centred design approach. Detailed results can be found in Appendix D. The layout diagram is shown in Figure 4.12, the interaction state diagram is shown in Figure 4.13, and an illustration of the interaction is shown in Figure 4.14.

One-step Zoom We adopt the interaction presented in [2], named DTZ (Dual-to-Zoom), which uses one button to zoom into a point of interest, defined by the location of the point of gaze, and another button to zoom out. We dedicate a third button for panning within the zoomed image, as will

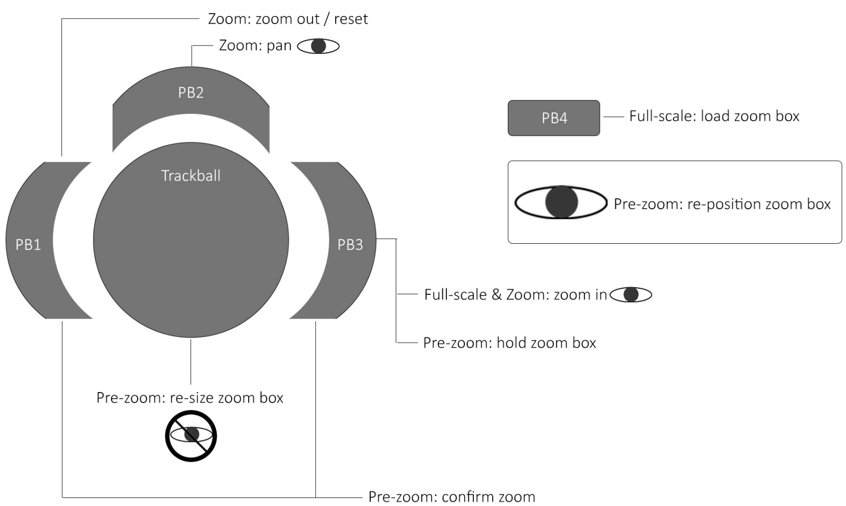


Figure 4.12: The Interface Layout of Design Alternative 1: the active input elements are the trackball, push button 1, push button 2, push button 3, and push button 4.

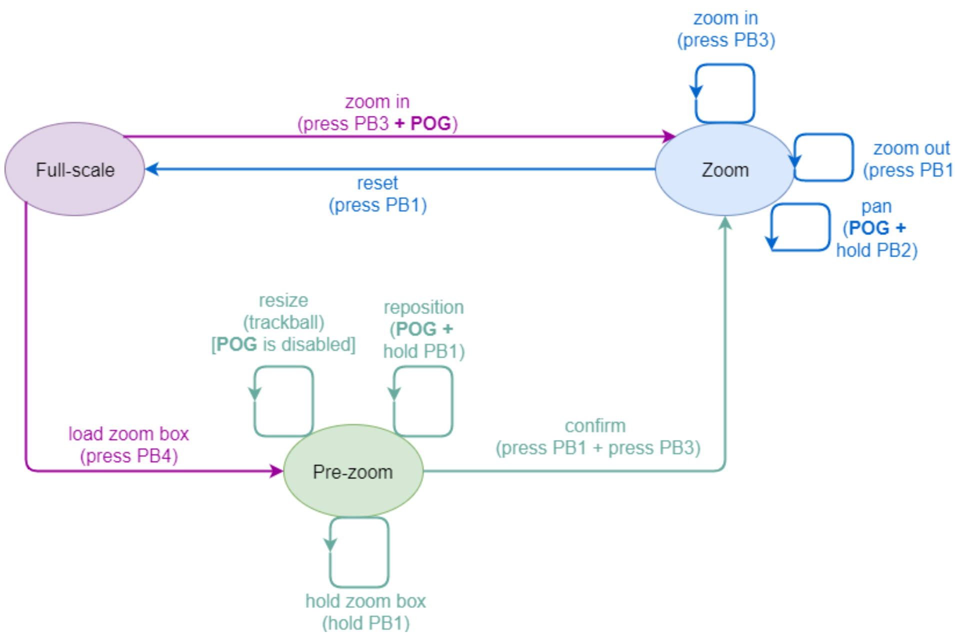


Figure 4.13: The Interaction State Diagram of Design Alternative 1: the total number of states are reduced by omitting the sub-states of Zoom and Pre-zoom.

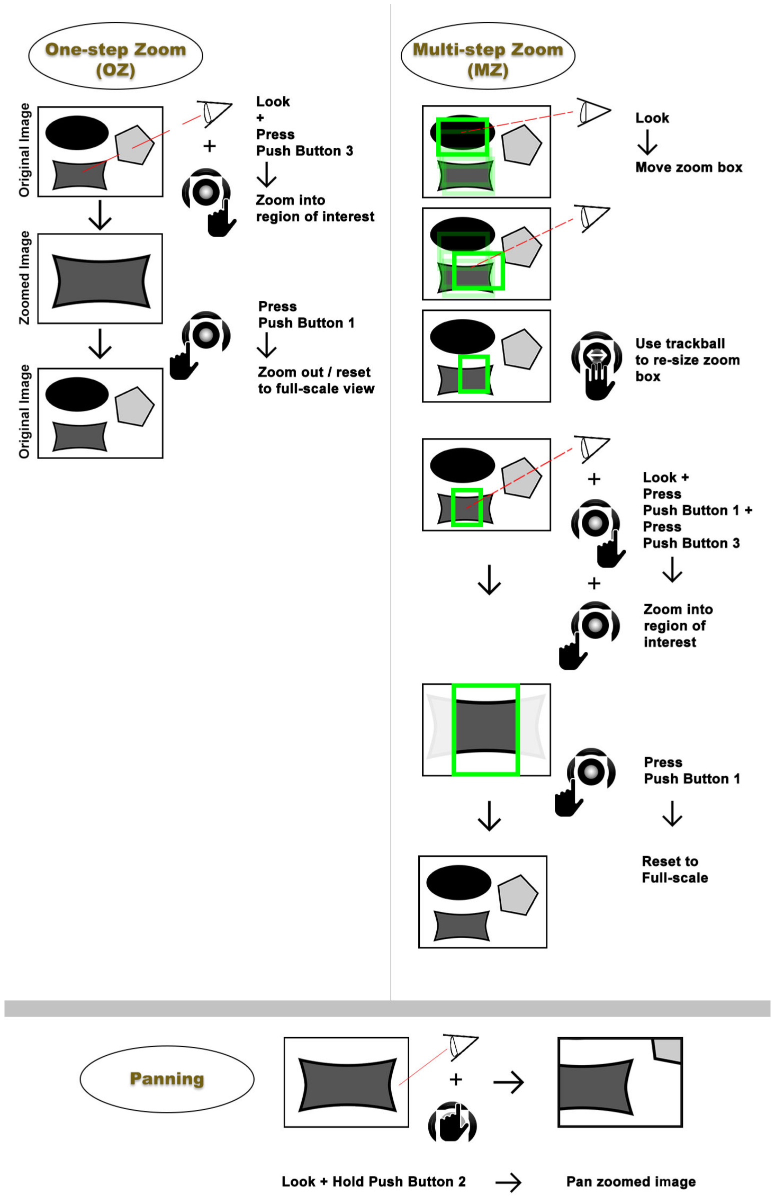


Figure 4.14: An Illustrated Interaction of One-step Zoom, Multi-step Zoom, and Panning of Design Alternative 1.

be detailed later. In our design, the first zoom action (120%) is performed based on the point of gaze. The consecutive "zoom-in" actions zoom in to the centre of the visible frame by a factor of 30%. "Zoom-out" always backtracks with every button click until the original image is restored.

Limitation Although this design alternative proved to have potential in the earlier work investigated [2], adopting it in ultrasound machines will require extra hardware (push buttons) that are dedicated for the OZ function. Given that we aim to integrate our gaze-supported interaction design with the currently-available premium to high-end ultrasound machine interfaces used in diagnostic sonography, we decide to explore a different design for One-step Zoom that adopts the same input devices used in the targeted ultrasound machines.

Multi-step Zoom

Repositioning the zoom box Once the user enables the zoom mode, a zoom box appears on the image and latches to the user's eye gaze. A simple averaging filter is applied, as shown in Figure 4.18, to reduce the jittery effect of rapid eye movements. Furthermore, the opacity of the box intensifies as the user gazes longer into a particular region and goes transparent again as the user rapidly looks away.

Resizing the zoom box The user can freely adjust the size of the zoom box through the trackball. Once the user starts resizing the box, the box stops following the eye gaze movement to allow for precision. In other words, the zoom box is not latched to eye gaze as long as the user is actively changing the size of the box. The user also has the option to press a button to manually "hold" the box in place.

Confirming the zoom area Finally, once the dimensions of the box are set and the location of the area of interest is held with a button, the user simultaneously presses another button to confirm the zoom action. The same zoom button is pressed again to zoom out and restore the original image.

Limitation Through testing with end-users, we find that the combination of button presses required for confirming the zoom function is cumbersome to learn and perform, let alone to perform repetitively every time the sonographer zooms into a target. Latching the zoom box to eye movement without a simultaneous button press caused a lot of distractions, as reported by sonographers and further detailed in Appendix D on the user study from the first iteration.

One of the alternative interface designs we thought of, which combines the visual feedback idea from the zoom box, but eliminates its distracting large shape, is replacing the zoom box with a cross-hair or a point that moves along with the eye. However, in addition to the previously discussed positive feedback loop issue, when the idea was presented to some of the end users, they pointed out that it might not be effective since some targets might be of an irregular shape and do not have a "centre" where the pointer should align with. This will cause additional cognitive load to the sonographer that will lead them to try and find the centre of a certain target and align the pointer with it before zooming into it.

Pan Regions For both OZ and MZ, once in the zoom state, pan regions are defined at the boundaries of a zoomed image, as shown in Figure 4.19. If the point of gaze falls within one of these regions while simultaneously holding the pan button, the image scrolls in that direction. This approach has been deployed in earlier gaze-supported zoom interactions, such as that presented in [2].

4.6.2 Alternative 2

This design alternative was explored in the second iteration of this user-centred design approach. We decide not to proceed with testing the interaction design with end users due to the limitations we realize through analyzing the interaction.

This interaction design alternative, shown in the input layout diagram in Figure 4.15, the state diagram in Figure 4.16, and the interaction diagram in Figure 4.17 greatly simplifies the interaction by reducing the total number of states the system transitions between as Zoom and Pre-zoom do not have sub-states anymore (compared to the original design in Figure 4.7). This is achieved through completely delegating all the position-related 2D input functions to the eye gaze input. Thus, panning the image in the Zoom state is simply achieved through holding a button and looking at the edges of the image. Similarly, moving the zoom box in the Pre-zoom state is simply achieved through holding a button and looking at the intended area.

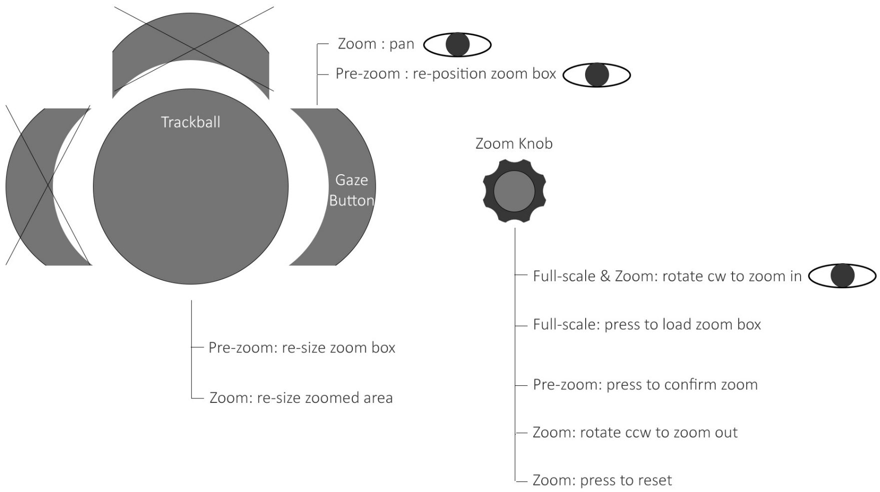


Figure 4.15: The Interface Layout of Design Alternative 2: the active input elements are the trackball, the gaze button, and the clickable zoom knob.

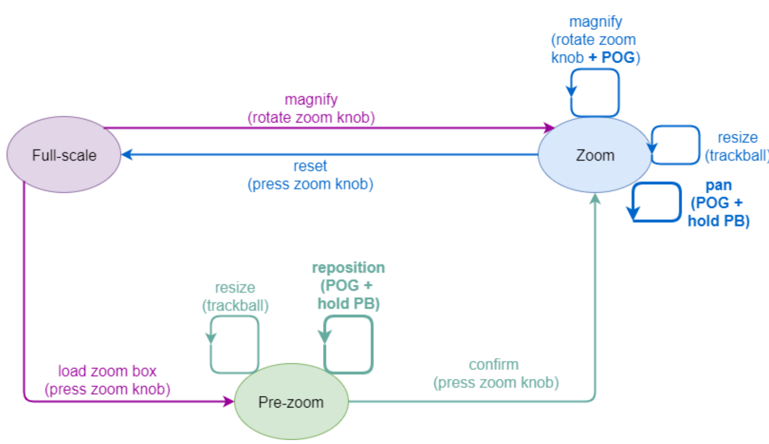


Figure 4.16: The Interaction State Diagram of Design Alternative 2: the combination of inputs is reduced compared to alternative 1, displayed in Figure 4.13, for some functions.

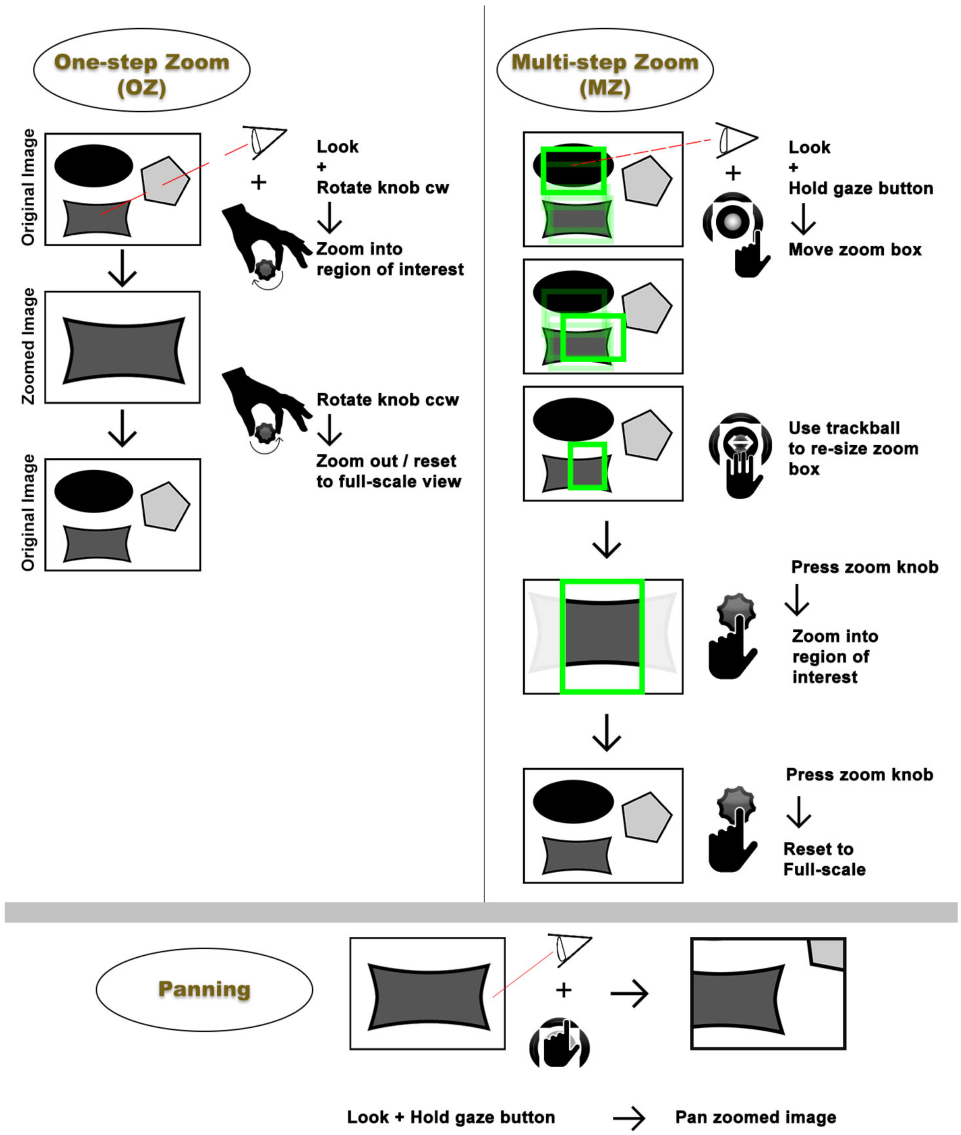


Figure 4.17: An Illustrated Interaction of One-step Zoom, Multi-step Zoom, and Panning in Design Alternative 2.

Limitations Theoretically, the design holds potential in simplifying the interaction and will inevitably reduce the repetitive interaction, especially by eliminating the need to toggle between sub-states. However, through pilot testing of the interaction, we realize that it is not feasible due to the following challenges:

1. Zoom state panning: looking at the edges of the image is counter-intuitive to what a sonographer (or any user performing a task requiring image inspection) does. When the user looks at the edge of the image and presses the pan button, the image moves to the left and therefore brings the object of interest inside the field of view. However, centring the object is not possible as the user must look back and forth between the edge of the image (to perform panning) and the target (to track its position).

2. Pre-zoom state box movement: although it works for long distance movements, however, as discussed in many previous eye gaze tracking studies and as observed in the results from the first iteration, fine accurate motions of objects on screen based on eye gaze input is not possible due to the eye's jittery movements.

4.7 Gaze-supported Features: Implementation Details

4.7.1 Filtering Gaze Data: Moving-average Filter With a Threshold

Although the gaze data generated by the eye gaze tracker is filtered, for the purposes of our control-based application, further filtering is recommended, as observed during our initial user study found in Appendix D. Thus, we adopt a simple moving-average filter, as used in earlier HCI studies related to gaze-supported interactions [33] [65]. Additionally, we activate the filter only when the gaze is moving in small distances. For larger distances, the filter is disabled. The window size and the rest of the constants used in this filter were determined through trial and error and a number of pilot tests. An overview of this algorithm is summarized in Figure 4.18. The overall interaction with the gaze-supported zoom interface, which uses this filter, is later evaluated in Chapters 5 and 6.

The following variables are used in the filtering algorithm. The selected values are determined based on trial and error.

- (x_{old}, y_{old}) are the coordinates of the previous unfiltered fixations.
- (x_{new}, y_{new}) are the coordinates of the current unfiltered fixation.
- d_list is the list of distances between successive fixations.
- d_win_size is the window size of d_list, we select a value of 10.
- $d\_threshold$ is a limit for the average distance between fixations. An average of d_list higher than this threshold indicates a high jump in eye movement.
- fix_list is the list of averaged (filtered) fixations within a distance less than $d\_threshold$.
- fix_win_size is the window size for the successive fixations within $d\_threshold$, we select a value of 100.

The algorithm below explains the filtering mechanism.

1. Block invalid fixations: fixations with an invalid flag (Validity = 0) or fixations outside the area of the screen are filtered.
2. Calculate the average distance between successive fixations for the past d_win_size fixations
   (a) Evaluate d
   $$d = \sqrt{(x_{old} - x_{new})^2 + (y_{old} - y_{new})^2} \tag{4.1}$$
   (b) append d to the list of fixation distances d_list
   (c) Evaluate $d\_average$
   $$d\_average = \frac{\sum d\_list}{d\_win\_size} \tag{4.2}$$
   (d) If $d\_average > d\_threshold$, clear fix_list and d_list
   (e) Else, append fixation $(x_{filtered}, y_{filtered})$ to fix_list with the following entry:
   $$x_{filtered} = \frac{\sum fix\_list_x}{size(fix\_list)} \tag{4.3}$$
   $$y_{filtered} = \frac{\sum fix\_list_y}{size(fix\_list)} \tag{4.4}$$

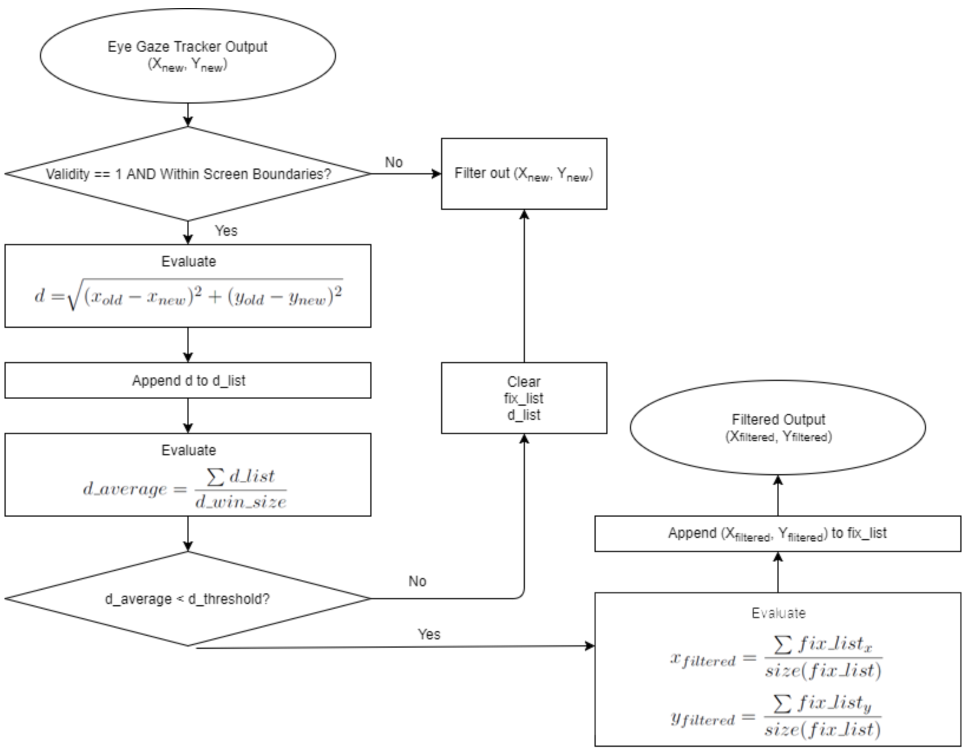


Figure 4.18: The Gaze Data Filtering Algorithm Used

4.7.2 Gaze-based Simultaneous Pan And Zoom

Provided the jittery nature of eye gaze data, zooming into where the user is looking will not yield accurate results: zooming enlarges the target and changes its central location causing the eye gaze to keep shifting as the user is gradually zooming in. This challenge has been acknowledged in earlier eye gaze tracking user interfaces research, such as the work investigated by Kumar [41], where he explains how the eye gaze error increases with increased zoom levels for interacting with maps:

"If the user was looking at point P, chances are that the eye tracker may think that the user is looking at the point P+ϵ, where ϵ is the error introduced by the eye tracker. Once the user initiated a zoom action, the map is magnified. Therefore, if the zoom factor is z, then the resulting error gets magnified to zϵ, which can be considerably larger than the original error."

To reduce this effect, we apply simultaneous zooming and panning, where the image zooms into where the user is looking, then pans to correct for the shifted target central position. A similar approach has been used in other gaze-supported control applications, such as implemented by [69]. We follow the same gaze-driven camera control algorithm in centring the target after the zoom action for a limit of 500 milliseconds.

The following variables are used in the simultaneous panning and zooming algorithm:

- POG is the filtered input point of gaze
- C is the visible image centre
- r_o is the radius around the centre
- $IM\_velocity\_max$ is the maximum velocity the image moves at
- D is the maximum distance in the visible image to the centre ($\frac{image\ diagonal}{2}$)

The following describes the simultaneous panning and zooming algorithm based on the work presented in [69].

1. Evaluate:

$$d = |\overline{POG\ C}| = \sqrt{(POG_x - C_x)^2 + POG_y - C_y)^2} \tag{4.5}$$

$$\theta = atan2(|POG_y - C_y|, |POG_x - C_y|) \tag{4.6}$$

2. If $d < r_o$: image remains in the same position.

3. Else:

   (a) Evaluate:

$$FG = \frac{d}{D} \tag{4.7}$$

$$IM\_angle\_current = \theta \tag{4.8}$$

$$IM\_velocity\_current = FG * IM\_velocity\_max \tag{4.9}$$

4.7.3 Gaze-based Panning Based on Pan Areas

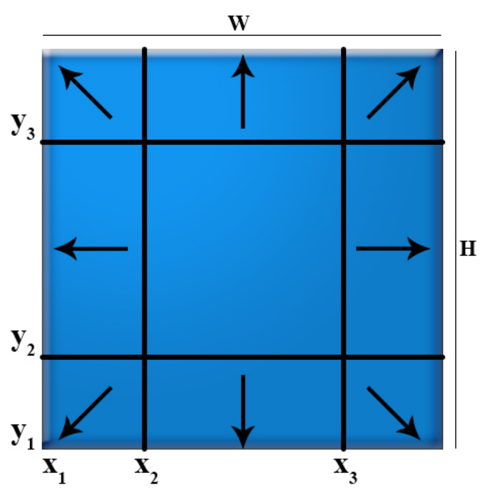


Figure 4.19: Gaze-supported Interaction Pan Areas Located at the Edges of the Image (8 Areas)

Another panning algorithm used in our design is moving the image based on the area where the point of gaze falls. As described earlier, this technique has been used in other gaze-supported pan and zoom work, such as [2] and [46]. Figure 4.19 shows the pan areas as the blocks enclosing the arrows. Each arrow points in the direction of image movement when the point of gaze falls in that block. If the point of gaze falls in any of these areas, the image translates with a speed equal to *IM_velocity_max* (refer to the previous section).

The ratios in Figure 4.19 are as follows: $x_1 = 0$, $x_2 = 0.25W$, $x_3 = 0.75W$, $y_1 = 0$, $y_2 = 0.25H$, $y_3 = 0.75H$.

It is also important to note that $IM\_velocity\_max$ is adjusted according to the zoom level for both panning algorithms based on a third-degree polynomial: $speed\ ratio = ax^3 + bx^2 + cx + d$, where x is the zoom ratio.

4.7.4 Mechanism of Gaze-supported One-step Zoom

For OZ, a combination of both panning algorithms (simultaneous pan and zoom and panning based on pan areas) is used. After zooming into the location of point of gaze, the image automatically pans for 500 milliseconds based on the following criteria: if the point of gaze does not fall in any of the pan areas, pan using simultaneous pan and zoom, otherwise, use the pan areas algorithm to pan.

This only applies for zooming in. Zooming out simply zooms out from centre regardless of the location of the point of gaze.

4.7.5 Mechanism of Gaze-supported Multi-step Zoom

For MZ, the only gaze-supported feature is an explicit action of moving the zoom box in the Pre-zoom state based on the location of the point of gaze when the gaze button is pressed and held. The same gaze filtering applied for OZ is also applied for MZ to reduce the zoom box's jittery effect due to the jittery movement of the eye gaze.

4.8 Custom Hardware Interface Implementation

Given that we are only investigating one class of functions, zoom and pan, and that we do not have access to the developer API of the machines observed in hospitals and used for routine diagnostic scans, we design and implement a custom-made ultrasound machine interface closely matching the interface design of the targeted ultrasound machines and stream ultrasound data from another class of ultrasound machines that we have access to in our lab.

The custom hardware interface, shown in Figure 4.20, is created to mitigate having the results be specific to the Ultrasonix Touch, the available machine in our lab with developer-level access. Our hardware interface includes only the relevant ultrasound functions for the evaluation presented in Chapter 6. The controls box is operated with an Arduino Mega that is connected to the main computer. The controls box has 5 rotary encoders:

one of them is for the zoom level and the rest are built to control the most commonly-used ultrasound image settings: gain, depth, frequency and focus. This interface is used as a manual input for both user studies described in Chapters 5 and 6. Ultrasound Images are transferred in real time with no observable delay through a TCP/IP connection.

Figure 4.20: The Custom-made Manual Controls Box Used for Both User Studies Described in Chapter 5 and 6.

4.9 Evaluation of the Presented Designs

We test the proposed interaction design and eye gaze-supported features through two user studies: the first is presented in Chapter 5 through a game-based user study that tests the system outside the environment of sonography, and the second is presented in the Chapter 6 through a clinical-based user study performed by sonographers.

4.10 Conclusion

We present a state analysis of the traditional manual-based One-step Zoom, Multi-step Zoom and panning in ultrasound machines. For the gaze-supported design, we follow the assumptions that all zoom functions acquire the magnified image with no improved quality and panning is enabled for both zoom approaches. We present the interaction state diagram of the final proposed design and the earlier designs explored and the interface layout diagrams and illustrated interactions. Finally, we present the implementation details of the gaze-based features: gaze data filtering with a thresholded moving-average filter, gaze-based simultaneous pan and zoom and gaze-based panning based on pan areas. We also explain how these gaze-based features are used in One-step and Multi-step Zoom.

Chapter 5

Context-free User Study: Interactive Game

We present in this chapter a game-based user study that tests the system outside the environment of sonography. From this study, we collect preliminary results to form an initial understanding of the gaze-supported interaction to help us better design the context-focused clinical-based user study, presented in Chapter 6, and anticipate the potential advantages and risks of the gaze-supported interaction before testing with end users.

This approach has the benefit of performing *basic interaction testing*: tasks involved in sonography have several factors external to the basic zoom interaction that could influence the performance of the user, such as managing the bimanual interaction, probe positioning, image analysis and communication with the patient. Running a study that is solely focused on zoomed targets acquisition will help us understand the interaction of the sonographers with the system better as it will help isolate effects due to the gaze-supported interface design from effects due to sonography tasks.

This user study is divided into two separate experiments: One-step Zoom Experiment (OZE) and Multi-step Zoom Experiment (MZE). We present the game design and structure, the experiment design, tools and results. We conclude by explaining the related aspects of the results of this user study to the next user study.

5.1 Goal

This chapter presents an exploratory study designed to get an overview of the performance of the gaze-supported features by qualitatively evaluating the cognitive load and manual repetitiveness and quantitatively evaluating time on task and accuracy. It compares the designed gaze-supported system for ultrasound machines to the manual-based system by keeping the same basic interactions, software and hardware interfaces of the zoom interaction designed for ultrasound machines and changing the target application to an area that is more simpler and suitable for a more general set of users. By referring to the work presented in [15], which describes the fidelity of simulations into two components: physical fidelity, which is composed of equipment fidelity and environment fidelity, and psychological fidelity, which is composed of task fidelity and functional fidelity, the prototype tested in this user study implements the equipment and functional fidelity dimensions. The equipment being used in this user study is the same, or resembles as close as possible, the equipment of an ultrasound machine interface. The functional fidelity is the same as we are testing the same gaze-supported zoom functions that we developed for use by sonographers in ultrasound machines.

This game study is designed to resemble as close as possible the targets and the environment of the clinical user study we present in Chapter 6, but with higher focus on the system interaction. Therefore, the tasks in this user study are not typical ultrasound imaging tasks, but resemble the same interaction techniques in the following aspects:

1. **Imaging Task.** A sonographer's task is typically to capture an image of the target of interest by showing the organ as clearly as possible and as large as possible within the image frame. In this study, the task of the participant is to locate the target of interest and zoom into it until its size takes up at least 80% of the image frame.

2. **Right-handed/Left-handed Setup.** In sonography (in most types of ultrasound exams), all sonographers are trained to hold the probe and scan with their right hand and use the manual controls with their left hand, even if they are left-handed. In our user study, we follow the same approach by restricting the interaction with the image controls to the left hand, regardless of the participant being right-handed, left-handed or ambidextrous.

3. **Targets.** Our clinical user study, as described in the next chapter, is designed for sonographers to perform a Common Bile Duct (CBD) scan using our proposed interaction designs. The CBD is typically only one target within the visible range of an ultrasound image. The tasks in this game user study are designed so that one target at a time shows up for the participant to zoom into and capture.

4. **Continuous Motion and Disappearance of Targets.** One of the challenges in sonography is finding the target, fixating the probe at the correct position and holding it to capture the image at the right

time as the patient is breathing. As the patient inhales and exhales, the image is distorted by fading in and out, as the organ being imaged enters or exits the ultrasound imaging plane, or by having the target move around as the internal organs are in constant motion. This adds a certain level of difficulty to the image zoom and capture task, so we design the game targets to have similar properties, to some extent, by disappearing periodically.

5.2 Experiment Design

Based on a few pilot tests of the designed interactions, we find that training a user, with no background in ultrasound interfaces, on both zoom techniques is too overwhelming for a one-hour long study. We observed that users often mixed up the two interactions, which affected the results as users did not have sufficient training. Therefore, we decide to separate this user study into two experiments to test the interactions independently: the One-step Zoom Experiment (OZE) and the Multi-step Zoom (MZE). The structure, design and tasks for both experiments are the same. The subtle differences will be explained in the following sub-sections of this chapter.

This user study is approved by the Behavioural Research Ethics Board at the University of British Columbia, under UBC CREB number H15-02969.

5.2.1 Game Design

The designed game is similar in concept to the classic arcade game of “Space Invaders”, as the main task of the player is to target and shoot the alien invaders. Similar art and sound effects were adopted in this game [18]. This similarity provides participants with more connection with the game as most of them are familiar with it. The following are the main differences between this game and the classic arcade game:

1. The player is provided with one alien (or group of aliens) and is required to eliminate them one at a time, instead of gradually eliminating multiple targets.

2. We adopt the “first-person” view instead of the “third-person” view implemented in the original game.

3. The game interface is re-designed to focus on the elements of the interaction being evaluated.

Game Software Interface

Figure 5.1 shows the software interface of the game. The "*Space Battle Area*" is where the targets show up to the participant, such as the green alien at the top. The "*Timer*" is a progress bar that fills up as more time is elapsed during the task. The "*Remaining Lives*" section displays the number of lives remaining before the game is over. Each life is represented as a blue heart. Each time a task times out, one life is lost. When all lives are lost, the game ends and the screen displays "*Game Over*". The "*Level Progress*" bar fills up as the player destroys more aliens during the game. Once the progress bar fills up completely, the player moves to the next level in the game. The "*Current Level*" box displays the level of the gamer in the game, as will be explained later in this section. The "*Eye Gaze Enabled/Disabled*" icon shows up when the system is actively using the user's eye gaze for some control input. For OZE, this icon is always displayed during the gaze-supported session. For MZE, this icon is displayed only when the user presses and holds the gaze button to move the zoom box. The "*Context View*" box is a display of the overall view, which shows the position of the zoom box after the user zooms in, as shown in Figure 5.2. The area for "*Trackball Function*" is used only with MZE. It shows a different icon based on the function of the trackball, whether it is for repositioning, as shown in Figure 5.3, or for resizing, as shown in Figure 5.4.

Game Hardware Interface

Figure 4.20 shows the controls box used in this user study. It is the same controls box used later in the user study presented in Chapter 6. Out of all the hardware interface elements, only five are used in this user study: the clickable zoom knob, toggle button, eye gaze button, trackball and capture button. The calibration button is also used, but only by the researcher. For OZE, the zoom button press, toggle button and eye gaze button are disabled. For MZE, the zoom knob rotation is disabled. Like the ultrasound machine interface, and as explained in Chapter 4, the zoom knob rotation performs OZ, the zoom knob press loads the zoom box and the toggle button switches between resizing and repositioning the zoom box. The eye gaze button, when held while the zoom box is loaded, moves the zoom box around based on eye gaze location. The capture button is the equivalent of "shoot" in the game.

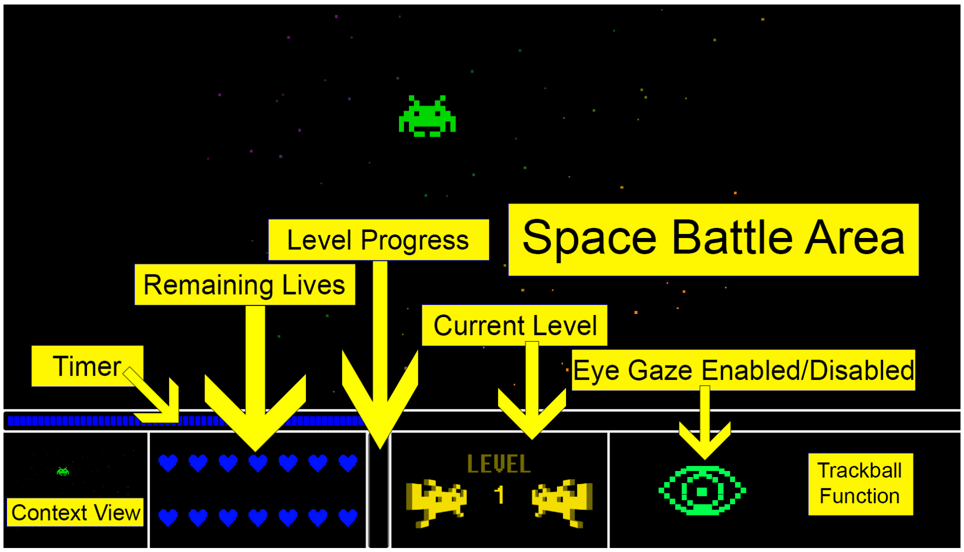


Figure 5.1: The Game Software Interface

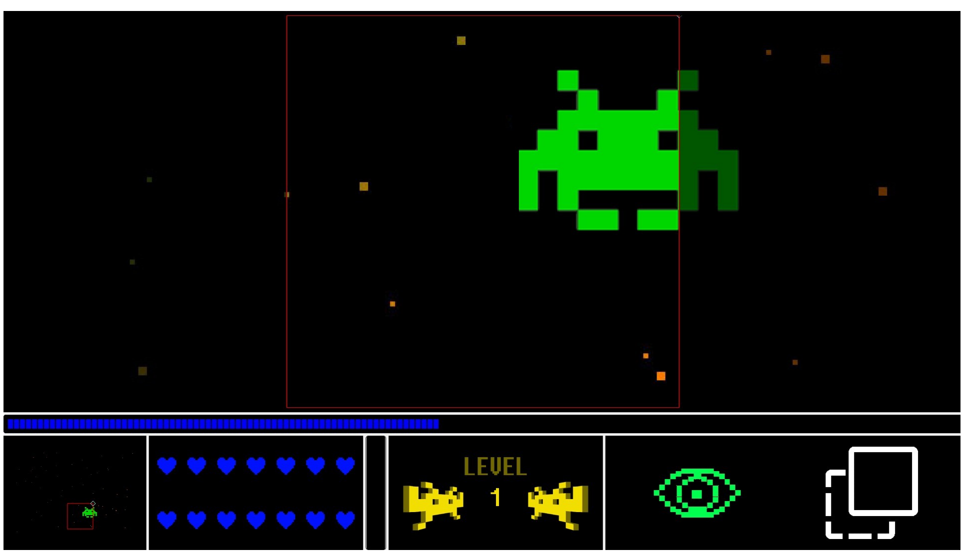


Figure 5.2: When zoomed into an alien, the context view shows the full-scale view with a box around the location of the zoom.

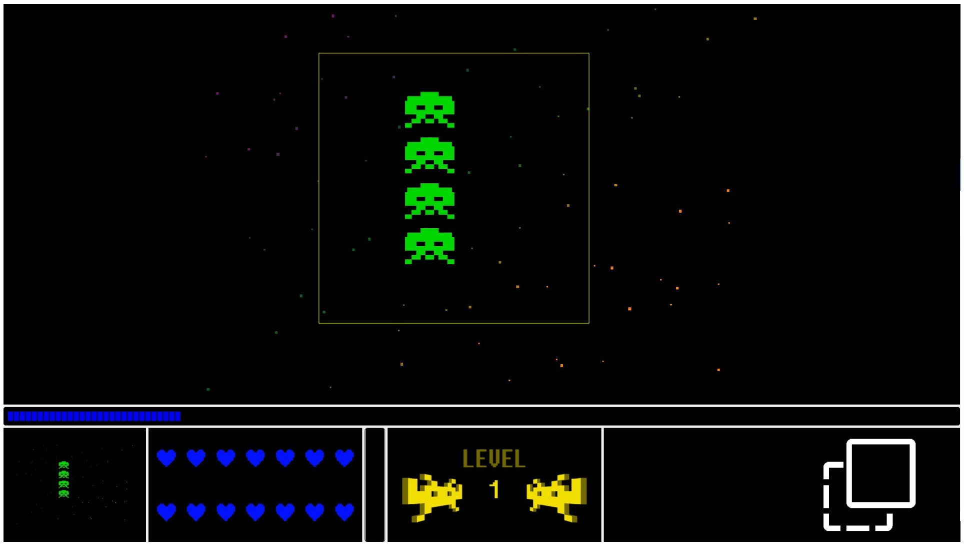


Figure 5.3: When the trackball is in reposition mode, the “reposition” icon is shown at the bottom right of the screen.

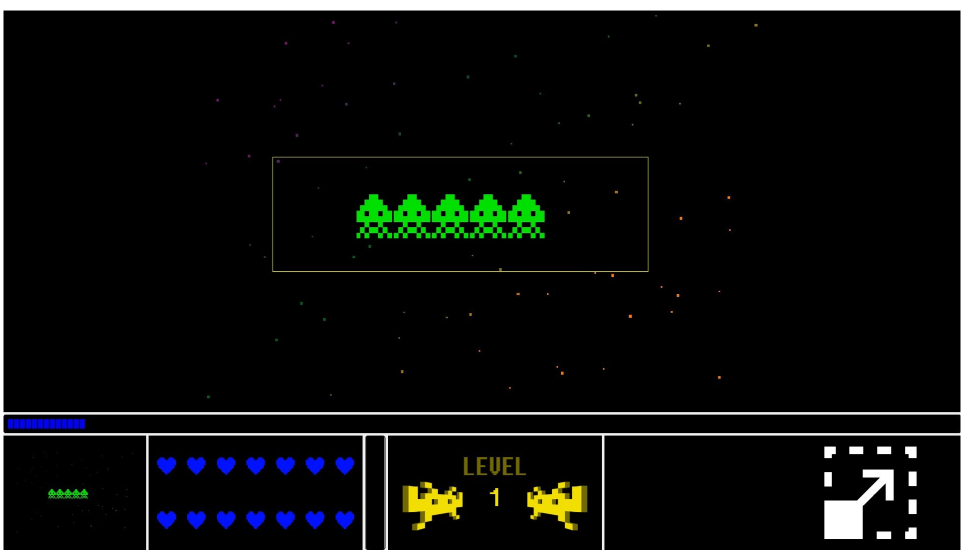


Figure 5.4: When the trackball is in resize mode, the “resize” icon is shown at the bottom right of the screen.

Gameplay Design

Targets At each level, aliens show up on screen one at a time. A player is required to shoot the target (for OZE, one alien, for MZE, a batch of aliens) before the time runs out. The colour of the aliens when they are first displayed is green, and they can be destroyed once they turn purple. Aliens turn purple only when:

1. They are fully within the zoomed view. Thus, if part of the alien is outside the view, it will turn green again.
2. They are at a specific zoom level, filling up 80% of the view. The participant is informed to change the zoom level until the alien turns purple. If it's too zoomed in or zoomed out, the alien turns green again.

For OZE, only one alien is the target, and for MZE, a batch of aliens are the target, as shown in Figure 5.5. Hitting the shoot button will destroy one or a batch of aliens altogether, if they are purple.

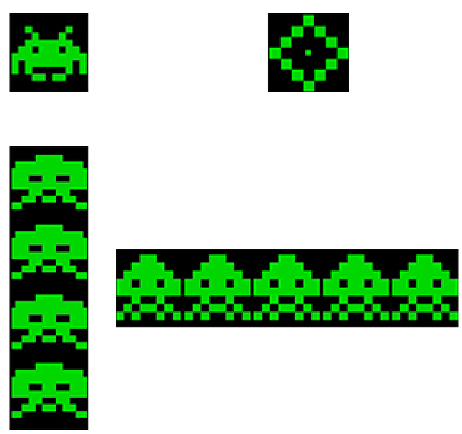


Figure 5.5: OZE Alien Targets (Top) and MZE Vertical and Horizontal Aliens Targets (Bottom)

As discussed earlier, to simulate difficulty found in sonography tasks, targets appear for 3 seconds and disappear for 1 second continuously in a periodic pattern. Once the alien(s) turn purple, they stop disappearing and start blinking to a lighter purple colour every half a second. Once an alien (or a batch of aliens) appears on some position on screen, it stays in the same position until destroyed.

OZE Targets The targets presented to the participant are randomly-generated with a maximum size of 300 x 300 pixels and a minimum of 100 x 100 pixels. The target position is also randomly-generated with a target distance from centre ranging from 2 pixels to 382 pixels. The targets are also randomly generated with an equal chance for the first alien shape and the second alien shape. Figure 5.6 shows the relationship between target size and distance from centre for all the generated shapes during the One-step Zoom experiment for all participants. As the shapes get larger, they grow closer to the centre, due to the boundaries of the screen.

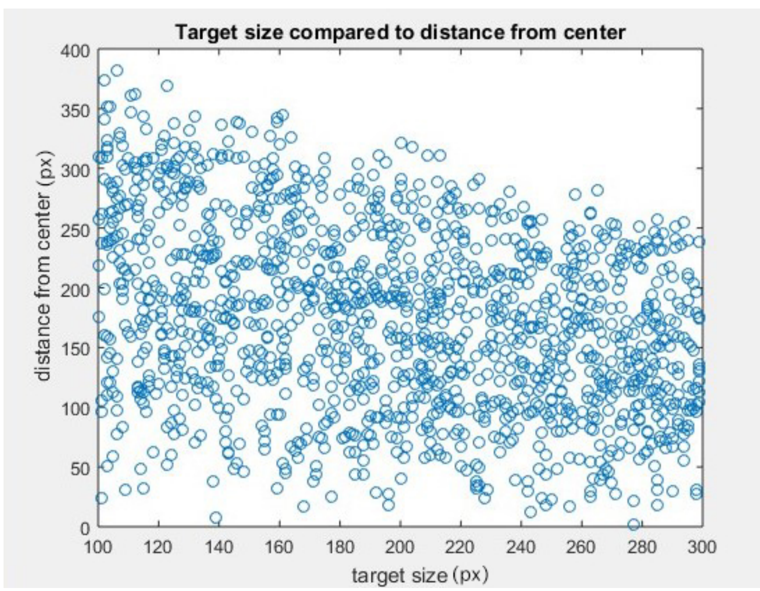


Figure 5.6: A graph showing the relationship between the generated targets' size and distance from centre during the One-step Zoom Experiment

MZE Targets The targets presented to the participant are randomly-generated with a maximum area of 44,289 pixels^2 a minimum area of 24,900 pixels^2, a maximum width/height of 400 pixels and a minimum width/height of 83 pixels. The target position is also randomly-generated with a target distance from centre ranging from 6 pixels to 350 pixels. The shapes are also randomly generated with an equal chance for a horizontal or a vertical order.

Scoring and Leveling Up Every time the player destroys a target (an alien in OZE or a batch of aliens in MZE), the level progress bar fills up by one point. A total of five consecutive destroyed targets will fill up the progress bar completely and will take the player up to the next level. If the player misses a target, all the progress made in the level will be lost and the player must restart the level. If all lives are lost, the player loses the game.

5.2.2 Setup and Structure

Figures 5.7 shows the setup of the context-free user study discussed in this chapter. The user study took place outside our labs, in a separate room reserved for running user studies at the University of British Columbia. The setup consists of two monitors connected to the same machine: one for displaying the game for the user, and the other for the researcher to observe the game in real time and the eye gaze tracker performance through Gazepoint Analysis [20]. The participant's monitor has the eye gaze tracker mounted at the bottom. The participant is provided with the controls box, headphones for the game sound effects and soundtracks and a seat with an adjustable height. The researcher is seated away from the participant, keeping minimal distraction. Participants were recruited through mailing lists. Participants who emailed back expressing their interest in participating in the user study were asked to provide their responses to the following eligibility questions:

1. Do you wear glasses?
2. Do you wear bifocal/gradual glasses (the ones for both far-sighted and short-sighted vision correction)?
3. Do you have any abnormal (whether diagnosed or undiagnosed) eye condition? (e.g. lazy eye after fatigue)
4. Do you have any left arm/hand/fingers injury? Do you have any pain associated with the movement of your left arm/hand/fingers for any reason?
5. Do you have any previous experience with operating ultrasound machines? (operating the machine itself, not being the patient)

In addition, a copy of the consent form was attached in the email and they were asked to choose a time within the period of their best performance during the day, if possible, as they will be performing tasks that require learning some new computer interaction techniques. If any of their answers to questions 2, 3, 4 or 5 were "yes", then they are disqualified from participating in the user study. If they answer question 1 with "yes", they are not disqualified, but they are recommended to wear contact lenses, if possible, as there had been some difficulties with eye gaze tracking with a few participants who wear glasses in previous pilot studies.

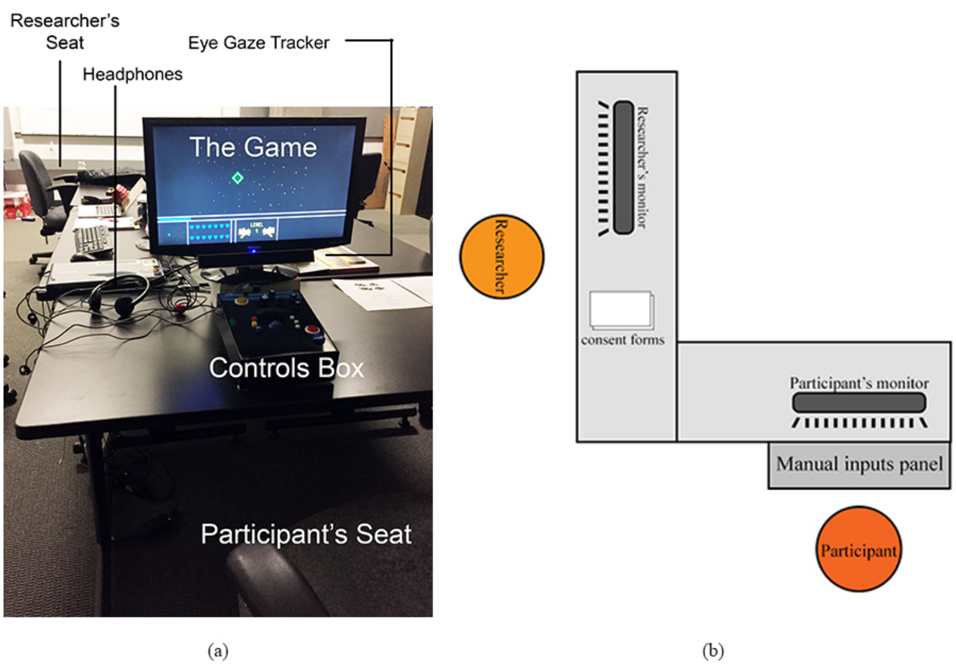


Figure 5.7: The Setup (Left) and Layout (Right) of the Context-free Game User Study

Table 5.1 contains the structure of the user study. As soon as the participant arrives, the participant signs a consent form, fills a demographics form (found in Appendix F), and receives a participation gift card reward. In addition, the eye gaze tracker is tested and calibrated with the gaze tracker's default 5-point calibration before the user study sessions to make sure there are no eye gaze detection issues. In addition, the headphones are also tested to make sure the volume level is comfortable for the participant. Eye gaze calibration is routinely tested before every sub-session that requires eye gaze input to ensure performance quality. A complete user study script is found in Appendix E. Demographics collected on the participants include their age range, being right-, left-handed or ambidextrous, their default setting of the touch pad scroll direction in their laptops, eye/vision conditions, their frequency of use of image editing or design software that require frequent zooming, and their situational level of mental tiredness. The only difference between sessions 1 and 2 is the input modality. This user study is counter-balanced, so half of the participants for each experiment had the gaze-supported interaction alternative for the first session and the manual-based interaction for the second session. The other half of the participants had the reverse order. Participants are permitted to interact with the researcher and ask questions only during the introduction, demo and post-experiment discussion sessions. During the training and recorded sub-sessions, the participant is instructed to completely focus on the game. The researcher ignores any remark or question made by the participant during the training and recorded sub-sessions, and interferes only in the case of a technical issue.

Training Sub-sessions

During the training sub-session, the participant keeps leveling up until they lose the game. There is no winning condition. The time limit per target for level 1 is always 20 seconds for OZE and 70 seconds for MZE. These values were determined through pilot tests. The time limit for each level that follows is equal to the average time elapsed for the five consecutive successful trials of the previous level. Therefore, the time limit for level 2 for participant 1 is different from participant 2, as the time limit is dependent on the user's performance. This level design aims to bring the participant to a level of saturated performance, where a participant cannot improve beyond their limits. This ensures there is are no carry-over effects from one session to the next, as the participant would have reached a maximum level of performance before the recorded sub-session for each input modality. There are 14 lives per level.

It is important to note that the mechanism of the training sub-session is not revealed to the participant before the user study, as it might cause the participant to intentionally slow down in the first few levels to keep winning, and thus go beyond the time limit of the user study and cause fatigue before the recorded sub-session. This was experienced during one of the test runs

Table 5.1: The Game User Study Procedure

Session	**Sub-session**	**Task(s)**
Introduction	The researcher introduces to the participant the user study procedure, the game software and hardware interface, winning conditions and tasks.	
Session 1	Demo	The researcher runs at least 3 tasks for the participant and assists, if needed.
	Training Tasks	The participant completes multiple levels of the game as detailed in the Training Sub-session.
	Recorded Tasks	The participant completes one level of the game, as detailed in the Recorded Sub-Session.
Break	The participant is given a break for a few minutes.	
Session 2	The same as session 1, using the other input modality alternative.	
Discussion	The researcher discusses a few topics with the participant, as detailed in the Post-experiment Discussion Session.	

in the lab before setting up the user study.

Recorded Sub-sessions

During the recorded sub-session, the participant plays only one level of the game. The finishing condition for the level is to destroy 14 targets. Contrary to the training sub-session, missing a target and running out of time does not cause losing all the progress made in the level, as the aim of the recorded sub-session is to measure the participant's accuracy and time limit and not to train the participant up to a specific level of performance. However, like the training sub-session, missing a target and running out of time causes the player to lose one life. Similarly, a player has 14 lives during the recorded sub-session level.

Post-Experiment Discussion Session

After the experiment, the researcher discusses with the participant the game experience. First, the participant is given space to provide their general feedback and questions about the game, if any. Second, the researcher requests the participant to list, if possible, three advantages and three disadvantages of each interaction modality. Finally, the researcher discusses the sources of frustration the participant listed down in the qualitative evaluation form, to understand them in depth. For MZE, the participant is further requested to explain the developed strategy for using the eye gaze feature, as it is enabled only upon the user's explicit button press and hold.

5.2.3 Apparatus

The software and hardware apparatus used in this user study is the same as the apparatus used in the clinical user study described in Chapter 6, minus the ultrasound machine and the real-time ultrasound image streaming.

The eye gaze tracker we use is the Gazepoint GP3 [20] Tracker, which has a visual angle accuracy of 0.5 to 1 degrees, a sampling rate of 60 Hz, and head box dimensions of 25 cm (horizontal movement), 11 cm (vertical movement) and ± 15 cm (depth movement). We use the default 5-point calibration for all our experiments. We also use the open standard API by Gazepoint [19] to implement the communication between the eye gaze tracker and the developed software.

The eye gaze tracker transfers in real time the participant's eye gaze data. The custom-made controls box is operated with an Arduino Mega that is connected to the main computer as well. Figure 4.20 shows the controls box, which has 5 rotary encoders. One of them is for the zoom level.

In addition, the trackball is surrounded by the toggle button (used in zoom modes) and the eye gaze button (to enable/disable eye gaze features, as explained in Chapter 4). Finally, the interface also includes a capture button to capture the alien.

The software interface is fully written in Python using PyQt4 and Pyqtgraph. The game design and graphics were all designed by the researcher using Adobe Photoshop, in addition to the graphics used from the original Space Invaders game [18].

The monitor dimensions are 57.2 cm (width) $\times$ 41.8 cm and the viewable size is 23.6 inches. The configured screen resolution for all experiments is 1920 $\times$ 1080 pixels.

5.3 Analysis Tools

Quantitatively, we look at the number of timeouts, the times elapsed per task and the learning curve per session. Qualitatively, we use evaluation tools, such as the NASA TLX [28].

5.3.1 Background on the qualitative tests used

NASA TLX

To evaluate the sources of task load, we adopt the NASA TLX evaluation and modify the description of each source of task load to fit the nature of the task performed by the participants. The description of each of these elements as presented to the participants can be found in Appendix G.

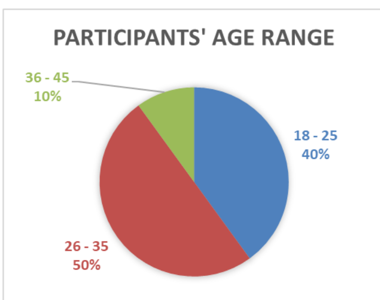
5.4 OZE Results

5.4.1 Demographics

A total of twelve participants were recruited for the study. Results from two participants were discarded due to eye gaze tracking difficulties. P3 had a noticeable case of amblyopia (lazy eye), which interfered with the eye detection by the eye gaze tracker. The reason for P7 is still unknown, like P3, the eye gaze tracker often stopped detecting the participant's eyes during the experiment. The rest of the statistical analysis is dependent only on valid results collected from ten participants: five males and five females. None of the participants wore glasses during the study, but only two participants wore contact lenses that did not interfere with eye gaze tracking. The age range of the participants is 18 to 45, as shown in Figure 5.8a. Two of the participants identify as left-handed, one ambidextrous and the rest identify as right-handed, as shown in Figure 5.8b. Right before the experiment, participants were asked to report on their situational level of tiredness. The majority of the participants reported as energetic, one fully energetic, one tired and two in-between, as shown in Figure 5.8c.

As the order of sessions is counter-balanced, Table 5.2 includes the order of sessions per participant. Order A means the participant completed the manual-based interaction session first then the gaze-supported interaction session second. Order B is the opposite to order A.

In terms of vision or eye conditions, P1 reported having a slight form of strabismus when tired, P2 and P12 had PRK vision correction surgeries and P5 had a Lasik vision correction surgery.



(a) Age Ranges of the OZE participants

(b) OZE Participants' Handedness

(c) Tiredness Level of OZE Participants

Figure 5.8: Demographics of OZE Participants

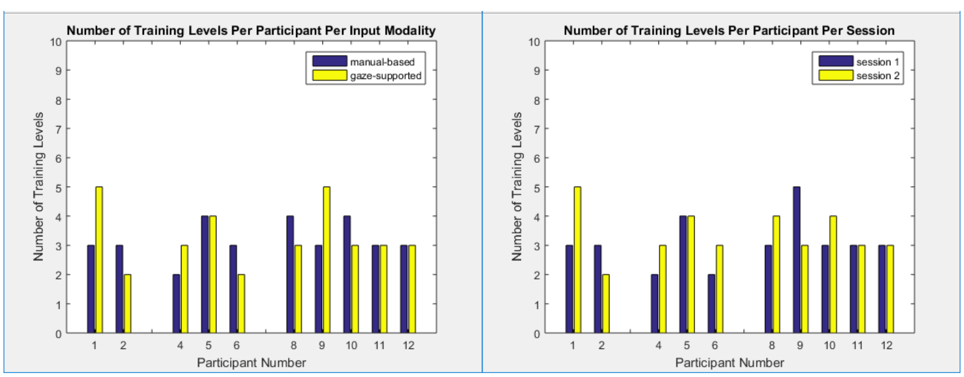
Table 5.2: OZE Participants' Session Order

Participant #	**Session Order**
1	A
2	A
3 (discarded)	A
4	A
5	A
6	B
7 (discarded)	B
8	B
9	B
10	B
11	B
12	A

5.4.2 Quantitative Evaluation

Training Sub-session

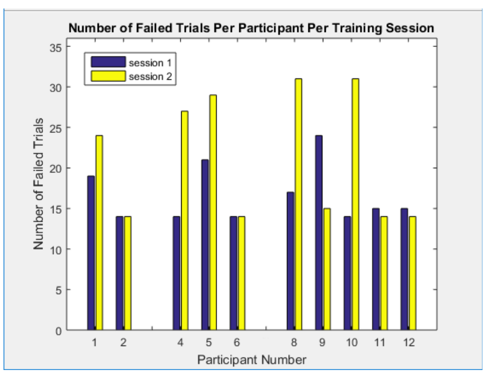
Number of Needed Levels First, we look at the number of training levels participants needed for each input modality, we find that there is no clear pattern, as some participants require more training for the gaze-supported interface and some require more training for the manual-based interface, as shown in Figure 5.9a.



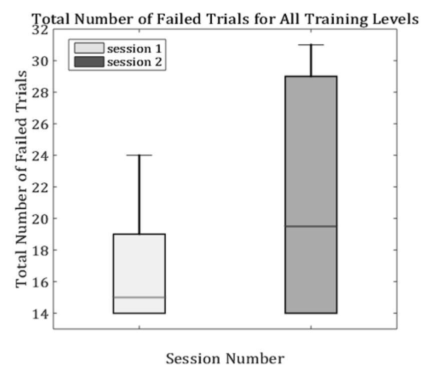
(a) Based on Session Name (b) Based on Session Number

Figure 5.9: Number of Training Levels per Participant per Session for OZE

Carry-over Effects Thus, we test for contamination effects between sessions; that the order by which the input modalities were introduced to the participant influenced the number of training levels required. In the first session, participants are introduced to the game and the manual controls for the first time, regardless of the input modality. This might pose an effect by increasing the amount of learning needed. However, by plotting the number of training levels needed per session, we fail to see any pattern again. This is shown in Figure 5.9b. We notice that some participants needed more training levels for the first session and some others needed more training levels for the second session. The average number of training levels needed for both sessions is 3 levels. By running Mann-Whitney U-test on the number of training levels needed per participant based on session order, we do not find a significant difference.



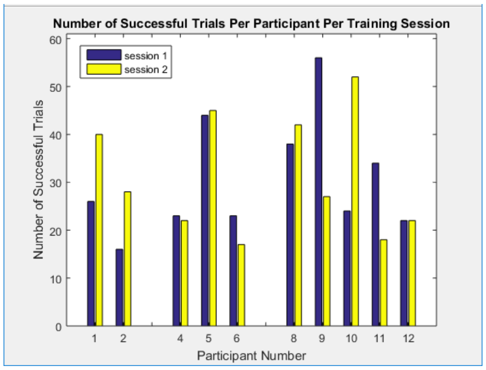
(a) Line Plot



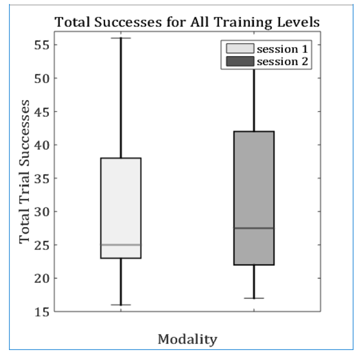
(b) Box Plot

Figure 5.10: Number of Failed Trials (Timeouts) per Training Session Number for OZE

To examine further, we count the number of fails and successes of trials, i.e. the number of times the task timed-out. We find that in most cases, session 2 had more or equal amount of failed trials than session 1, which could be due to the general fatigue of the participant. However, we cannot generalize, as it is not the case for participants 9, 11 and 12, as shown in Figures 5.10a and 5.10b. Moreover, the average number of successful training trials is not significantly different between sessions based on a Mann-Whitney U-test, as shown in the plots in Figures 5.11a and 5.11b. This shows that there is a minimal chance of contamination of data between the sessions and that participants require re-learning the interface when introduced with a new modality of input.



(a) Line Plot

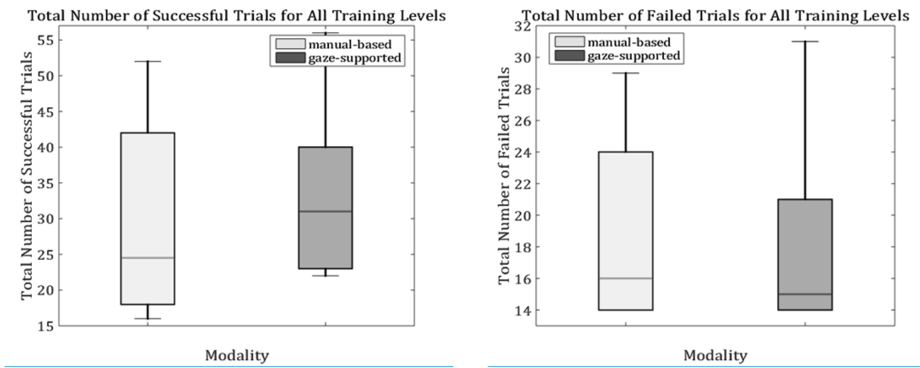


(b) Box Plot

Figure 5.11: Number of Successful Trials per Training Session Number for OZE

Accuracy Figures 5.12a and 5.12b show the total number of successes and fails for all training levels classified by input modality. We find that the gaze-supported interface learning session has a higher number of successes on average compared to the manual-based interface learning session. These results show potential for the gaze-supported interface in an increased accuracy, as the average successes are higher and the average fails are lower than the manual-based interaction. However, we also acknowledge the high standard deviation in both fails and successes.

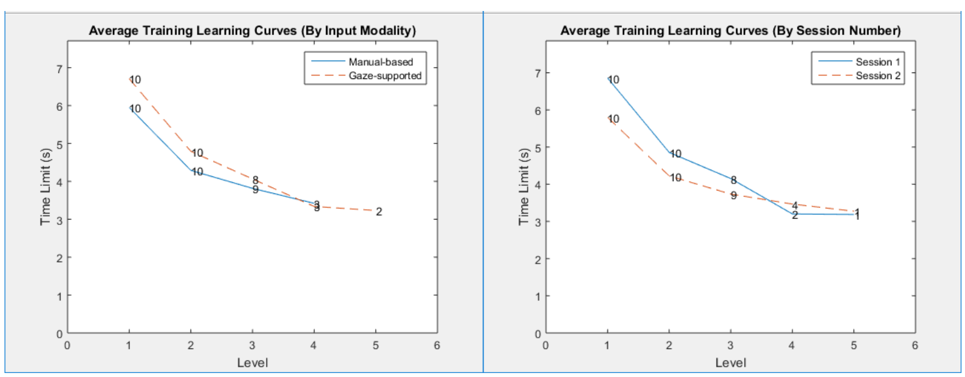
The Learning Curve As intended by the user study design, we observed a progressing learning curve, which appears to follow the power law of learning, for both input modalities across the training levels. Figures 5.13a and 5.13b show the average time limit per level for all participants, per input modality (Figure 5.13a) and by session order (Figure 5.13b). The number marked at each data point represents the number of participants that experienced this level. We observe that all participants passed level 1 and all participants started level 2. Only 8 participants passed level 2 and leveled up to level 3 using the gaze-supported interaction and 9 using the manual-based interaction. Only 2 participants passed level 4 and required a 5th level using the gaze-supported interaction. By looking at the same curve, but per session order (Figure 5.13b), we find that one participant from each group leveled up to level 5. In general, the time limit for session 1 is longer



(a) Successful Trials (b) Failed (Timed-out) Trials

Figure 5.12: Trials per Training Session Input Modality for OZE

than the time limit for session 2, however this cannot be generalized as the pattern was reversed at level 4.

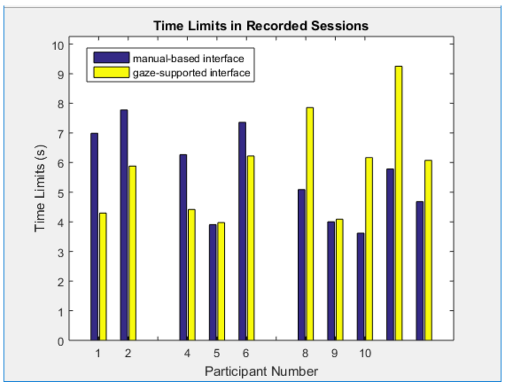


(a) By Session Input Modality (b) By Session Number

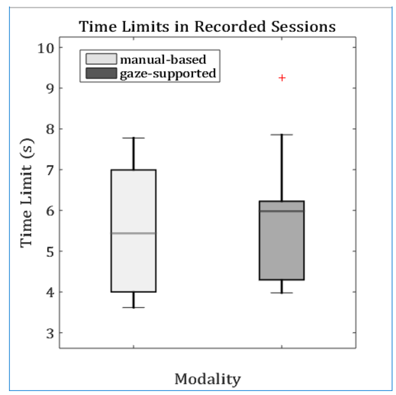
Figure 5.13: Participants' Learning Curve During OZE: The number marked at each data point represents the number of participants that experienced this level.

Recorded Sub-session

Time on Task We find that the average time limit for all participants in the recorded session is higher for the gaze-supported modality, as shown in Figure 5.14b. Figure 5.14a shows the time limits per participant for each modality. However, we also observe that the time limits are influenced by the user group, as participants in group A tended to have longer time on task during the first session (manual-based interaction), and participants in group B tended to have longer time on task during the second session (gaze-supported interaction). This shows that the time allocated for the user study of one hour, even with an extensive training session per interaction, might not be sufficient for conclusive results.



(a) Line Plot



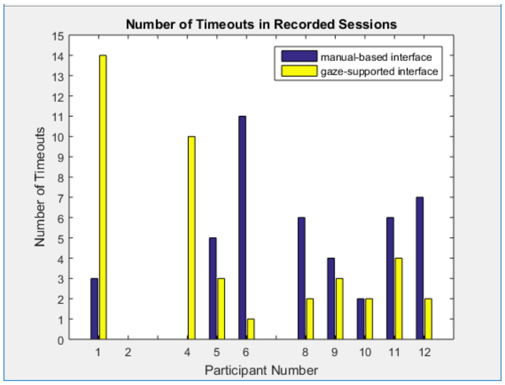
(b) Box Plot

Figure 5.14: Time Limits in the Recorded Sessions of OZE

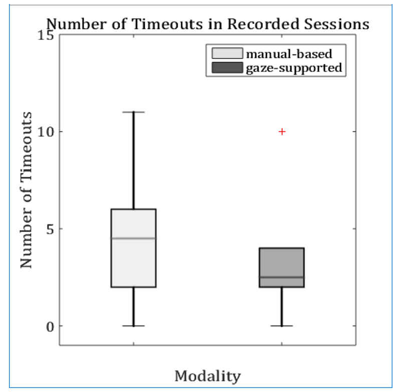
Accuracy Similar to the training sub-session, results from the recorded sub-session suggest that, on average, the number of timeouts for the manual-based interaction is higher than for the gaze-supported interaction. Results are shown in Figures 5.15a and 5.15b. These results also support the potential for a more accurate interaction using a gaze-supported interface.

5.4.3 Qualitative Evaluation

Through discussions, participants reported generally higher cognitive requirements for the gaze-supported interface, as will be explained in the Post-Experiment Discussions sub-section. However, by examining the recorded qualitative Task Load Index results, we find out that participants rated both input modalities very similarly. Figure 5.16a shows the overall TLX score for each interaction modality. We find that the mean of the manual-based group is slightly higher than the gaze-supported group. However, there is also large variation. Earlier work in eye gaze-supported interaction also



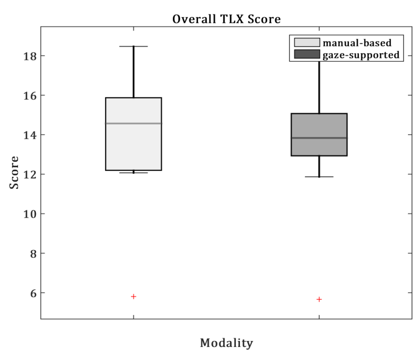
(a) Line Plot



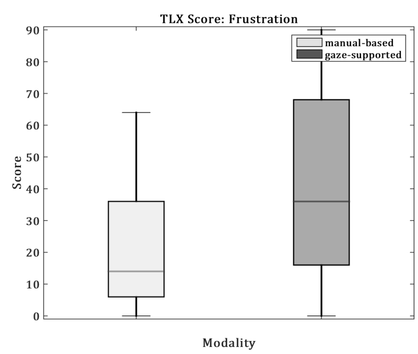
(b) Box Plot

Figure 5.15: Timeouts in the Recorded Sessions of OZE

used TLX as a qualitative evaluation for the sources of workload and found similar results. The work investigated by Klamka *et al.* [39], shows that TLX ratings did not have significant difference across the tested interaction alternatives. Similarly, Mental Demand and Frustration scored higher with the multi-modal gaze-supported interaction compared to the traditional manual-based interaction.



(a) Overall TLX Score



(b) TLX Frustration Score

Figure 5.16: Box Plots of the Qualitative Evaluation Results of OZE

Figure 5.17 shows a breakdown of the sources of task load. We find that both modalities are almost equal in all sources of task load, except for physical demand, where the manual-based system scores higher, and frustration, where the gaze-supported system scores higher.

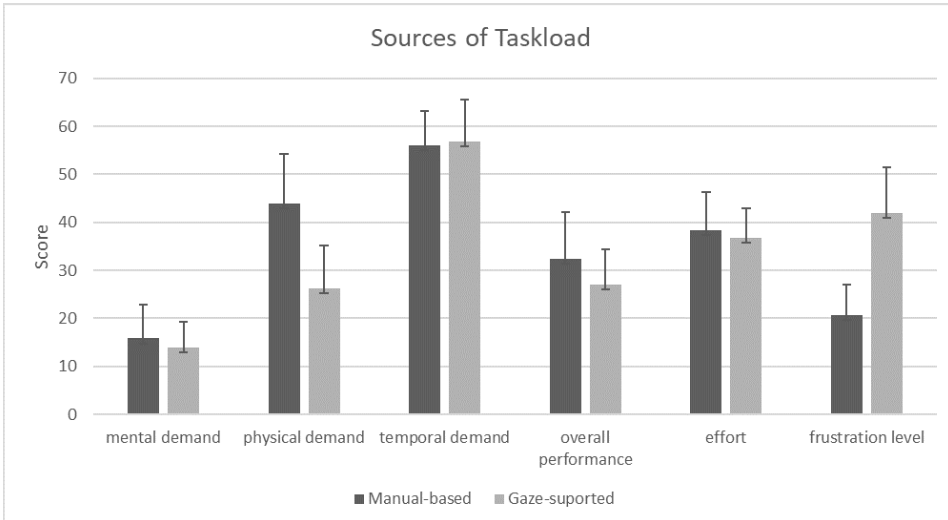


Figure 5.17: Sources of Task Load for OZE per Input Modality

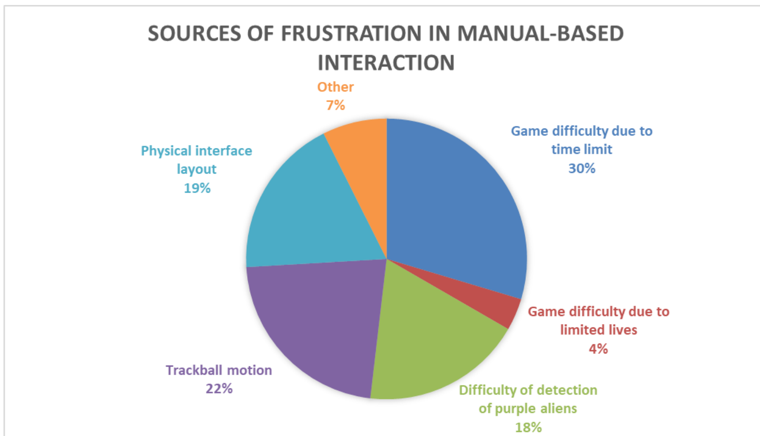
Mann-Whitney's U-tests were performed on all the sources of task load. Results show that there are no statistically significant differences between the two modalities in any of the sources of task load. However, we discuss below the small differences observed and how they triangulate with the post-experiment discussion with the participants.

Sources of Frustration

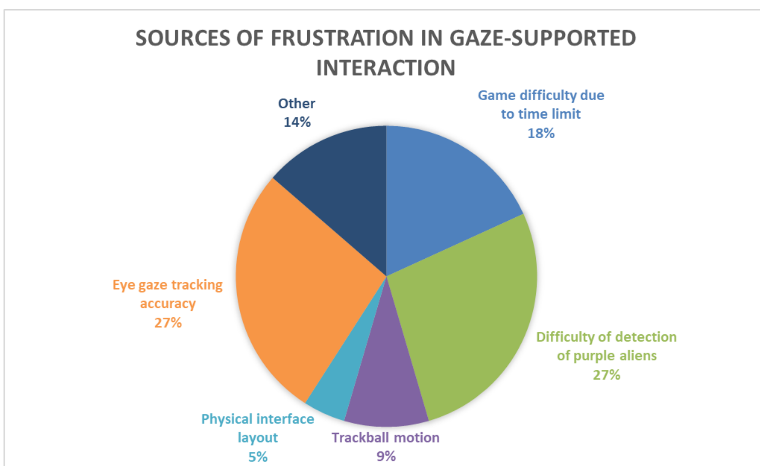
As shown in Figure 5.16b, there is difference between the manual-based system and the gaze-supported system in terms of the frustration level.

When asked to report the sources of frustration for each system, 30% of the participants reported the time limit as a source of frustration, followed by 22% of the participants reporting trackball motion, followed by 19% of the participants reporting the physical interface layout. The rest is shown in Figure 5.18a.

On the other hand, when asked about the sources of frustration in the gaze-supported system, 27% of the participants reported the eye tracking accuracy as a source of frustration. This is followed by 27% for the difficulty of detecting the purple aliens, followed by 18% for the game difficulty due to time limit. Details are shown in Figure 5.18b.



(a) Manual-based Interaction



(b) Gaze-supported Interaction

Figure 5.18: The Reported Sources of Frustration during OZE

5.4.4 Post-Experiment Discussions

During the discussion session, participants were asked for their general feedback on both systems, and the advantages and disadvantages of both. Given the challenging nature of the game, participants were also asked to elaborate further on their sources of frustration to be able to separate the interaction-related challenges and the gameplay-related challenges.

Advantages of the Gaze-supported Interaction

Reduced Physical Demand Participant 9 reported that using the manual-based interaction requires *"twice as much the physical effort"* as it requires repeatedly zooming and moving the trackball to pan. Participants also perceived higher speed with the gaze-supported system as they had *"less manual controls to worry about"*, as reported by P2. Some participants did not use the trackball during the gaze-supported tasks and preferred to zoom in an out with eye gaze-supported input repeatedly as an alternative to panning.

A Potential for a Higher Focus on Tasks P10 reported that it is more difficult to use the manual interface elements when focused on the target and prefers to use the eye gaze-supported system instead as it does not require a shift in attention from the screen to the manual controls as often. Some participants also reported that this reduction in physical interaction potentially introduces more focus on tasks on screen.

Disadvantages of the Gaze-supported Interaction

Unfamiliarity Participants described the manual-based system as consistent, reliable and overall simpler and easier to learn. This can be attributed to many reasons, including the familiarity of the participants with manual-based systems, in comparison to gaze-supported systems.

Higher Mental Demand The gaze-supported system was also described as more mentally-demanding as participants had to be always aware of their gaze location when the system is taking in their eye gaze input for control.

Another disadvantage of the gaze-supported system is the false perceived reliability of the system. P2 reported she expected during some tasks that looking at the target alone will zoom in and capture the target as the interaction is more automated than the conventional manual-based interaction. This overly-automated system could create some false expectations by the users.

Inaccuracies at Higher Zoom Levels Assume a target's centre covers 5 pixels at full scale, at twice the magnification, the target's centre will cover 10 pixels. This might not be an issue when using a manual system to zoom in, as the zoomed area is being constantly corrected by consecutive manual panning actions. However, when an eye gaze input is used, even if the user is instructed to look at the centre of the target to zoom in, the centre keeps getting larger and eventually there is no "real centre" anymore for the user to look at, which causes unwanted shifts in the zoomed area that will cause the target to be placed away from the centre of the magnified image. This shows that there are no improved accuracies at higher levels of zoom between the gaze-supported and manual-based interactions.

Unintended "*Midas Touch*" Effect The *Midas touch* effect in eye gaze tracking, as discussed by earlier research [32], refers to the situation when a user using eye gaze as a direct means of input accidentally issues a command wherever she looks, which is counter-intuitive to the typical function of the eyes. Having the manual input elements used during the zoom task not co-located, posed some challenges. In the manual-based system, the user is free to look away from the screen as she performs zoom actions, since the gaze is not an input to the location of zoom. The gaze-supported system, on the other hand, requires the user's gaze to lock onto the screen while zooming. In the case of gaze-supported interaction, the user does not have the freedom to look away and zoom in to the target, as the system has no input to where the target is. In case the user looks away while zooming using the gaze-supported system, one of the two following scenarios is bound to happen: either the system zooms into the default centre, which might not contain the target, or accidentally zooms into a lower portion of the screen the user has momentarily looked at as she moved her gaze from the screen to the controls and continues to zoom, which will also result in zooming into a false position.

Interaction-related Challenges

Physical Layout Design With the manual controls layout designed like a common ultrasound machine layout, many participants found it very challenging to perform gaming tasks, which is expected, given the nature of tasks in the game user study represent an extreme case of zooming into and capturing of targets. This extreme case is not encountered in routine sonography with the same amounts of time pressure. Participants, especially those with smaller hands, found it difficult to zoom by rotating the knob and moving the trackball at the same time. Thus, zooming and panning simultaneously was not feasible using the fully manual system.

Manual Movements at Higher Zoom Levels One issue reported by participants is the difficulty in applying manual movements at higher zoom levels, as the movement speed does not decrease based on zoom level. I.e. the panning speed is the same for all zoom levels, which makes a single pan action at zoom level of 400% move twice as many pixels as the image at a zoom level of 200%. Participant 4, for example, reported the usage of the trackball only if necessary at low zoom levels. To compensate for this issue at high zoom levels, during the gaze-supported session, he tended to use the zoom knob to zoom in and out to perform "panning with eye gaze".

Gameplay-related Challenges

Flickering Targets Like the reported challenges by sonographers of the difficulty keeping the target in the view due to its frequent movement and disappearance, participants in this game-based user study reported difficulty keeping the target (alien) in the filed of view due to its frequent disappearance. This challenge was meant to be experienced by the users as part of the experiment design.

Time pressure The perceived time pressure by the participants is due to several reasons. The adaptable time limit design in the learning session provided a sense of increasing game difficulty. Although participants were instructed to evaluate and report their feedback only on the 1-level recorded session, many of them perceived the second interaction, whether it is manual-based or gaze-supported, as more time-limited as they falsely-perceived, in most cases, the learning sub-session being shorter due to their familiarity with the game and the controls. For instance, participants who performed gaze-supported tasks in their second session reported a higher sense of increased speed of target acquisition due to eye gaze tracking compared to those who performed gaze-supported tasks in their first session. This false sense of speed is since they have already learned the manual interface and game challenges in the first session.

5.4.5 Discussion of Results

We analyze the results from the training levels and the recorded level and relate those findings to observations of user behaviour and post-experiment discussions with participants.

We find that there is no clear pattern in the amount of training levels required per participant between the gaze-supported or the manual-based interfaces. We also find no clear pattern when compared the number of required training levels per input modality order. This shows that there are no potential carry-over effects between the sessions and that the results per input modality are independent. However, some participants reported falsely-perceiving less required training for the second input modality, where it is gaze-supported or manual-based. We acknowledge an observed effect on time on task based on session orders during the recorded sub-sessions.

Although we find high standard deviation in terms of accuracy between the two input modalities during both training levels and the recorded level, the average number of timeouts for the gaze-supported input modality is lower than the number of timeouts for the manual-based input modality.

Participants reported higher frustration overall with the gaze-supported interaction, which could be attributed to the additional sources of cognitive processing introduced with this multi-modal interaction. They also reported higher physical demand associated with the manual-based interaction.

Advantages of the eye gaze-supported interaction, as mentioned by participants, include a higher sense of focus on the tasks at hand and a lower overall physical demand. Disadvantages include the general unfamiliarity with gaze-supported systems, higher mental demand and inaccuracies at higher zoom levels, in addition to required changes in the manual interface to better suit a multi-modal interaction.

The potential areas where a gaze-supported system might contribute in improving are a reduced physical interaction and a higher attention to the main tasks. However, with it comes a risk for other challenges, such as higher mental workload, and higher frustration due to multiple reasons including lack of familiarity with the multi-modal interaction. We examine these effects later within a clinical context in the user study presented in Chapter 6.

Task Load is Different for Each Interaction Participants reported experiencing task load in both systems. When asked to elaborate further, it was found that the type of task load is different for each system. For the gaze-supported system, the task load is due to high mental demand. While for the manual-based system, it was due to high demand of physical inputs coordination. Although gaze-supported systems alleviate the need for manual inputs coordination, as it requires less manual inputs, they introduce an increased mental demand requiring complete focus on the target on screen.

Eye Gaze Input Integration Requires a Change in the Physical Layout As evident by some of the challenges faced, to efficiently integrate an eye gaze tracking system into an existing manual-based interface, changes in the physical interface should be made. Physical interface elements used during an activity that requires eye gaze input should be designed to be co-located, which minimizes the need for the user to shift her attention from the screen to the manual inputs and thus minimize "unintended *Midas Touch*" effects.

5.5 MZE Results

5.5.1 Demographics

A total of nine participants were recruited for the study. Results from two participants were discarded: P6 due to eye gaze tracking difficulty and P9 due to an inconsistency in providing instructions to the participant with the rest of the sample, which directly affected the performance. The rest of the statistical analysis is dependent only on valid results collected from seven participants: four males and three females. Four of the participants wore prescription glasses during the study and the rest had no vision correction. The age range of the participants is 18 to 35. All the participants identify as right-handed. Right before the experiment, participants were asked to report on their situational level of tiredness. Most of the participants reported as energetic, one fully energetic, and one tired. Since the experiment order is counter-balanced, Table 5.3 shows the order for each participant, where order A is for manual-based session followed by the gaze-supported session, and order B is the opposite.

We did not proceed to recruit more participants for this experiment as we observe a pattern of carry-over effects between the two sessions per participant, as will be explained throughout this section. Therefore, we use the results from this experiment to report on the general observed user behaviour with the gaze-supported features with the carry-over effects limitations of this study.

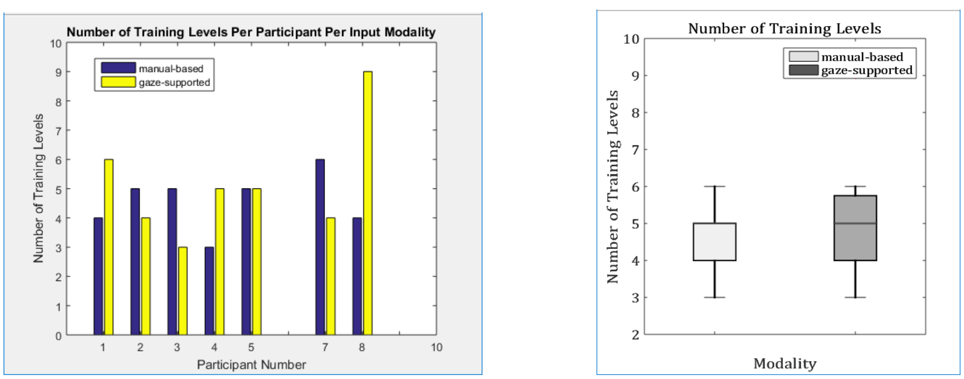
Table 5.3: MZE Participants' Session Order

Participant #	Session Order
1	A
2	B
3	A
4	B
5	A
6 (discarded)	B
7	B
8	A
9 (discarded)	A

5.5.2 Quantitative Evaluation

Training Sub-session

Like the results from OZE, we find that there is no clear pattern in terms of number of training levels between input modalities, as some participants require more training for the gaze-supported interface and some require more training for the manual-based interface, as shown in Figure 5.19a.

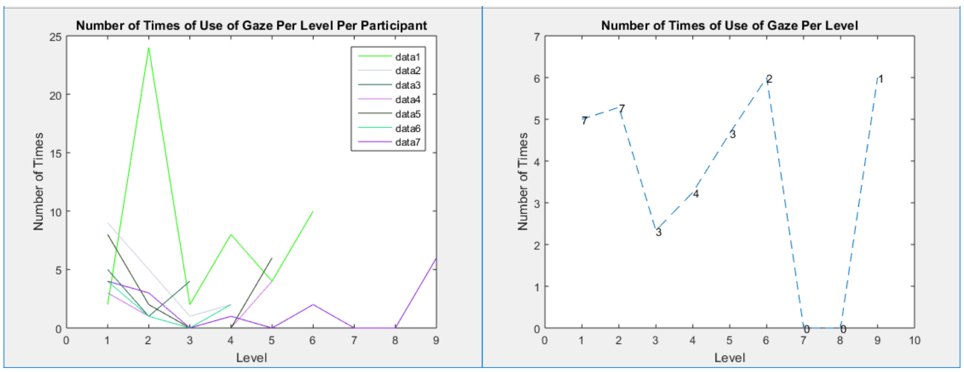


(a) Line Graph (b) Box Plot

Figure 5.19: Number of Training Levels per Session for MZE

The box plot in Figure 5.19b shows the number of training levels needed for participants per modality. We see that participants in general required more levels for the gaze-supported session compared to the manual-based session, although the difference is with a high standard deviation in both modalities.

We observe during the user study that participants' pattern of using the eye gaze input changed over the levels. Figure 5.20a shows the amount of usage of the gaze feature during the gaze-supported learning session. We find that, except for participants 1 and 4, participants generally reduced their use of the gaze feature as they level up. Figure 5.20b shows the average number of times the gaze feature was used per level for all participants. The number at each data point represents the number of averaged data points (participants) at that level. This shows that the gaze feature is found potentially inefficient with increased time pressure.



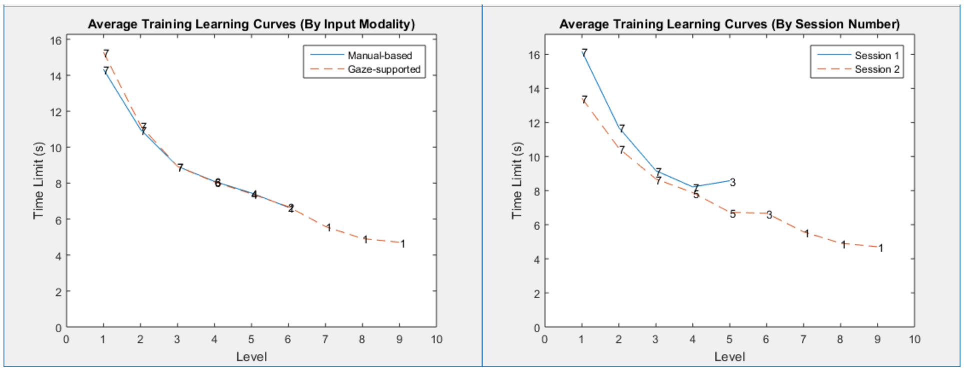
(a) Per Participant

(b) Averaged per Level: The number at each data point represents the number of averaged data points (the number of times the gaze feature was used) at that level.

Figure 5.20: Number of Times the Gaze Feature was Used During the Gaze-supported Session of MZE

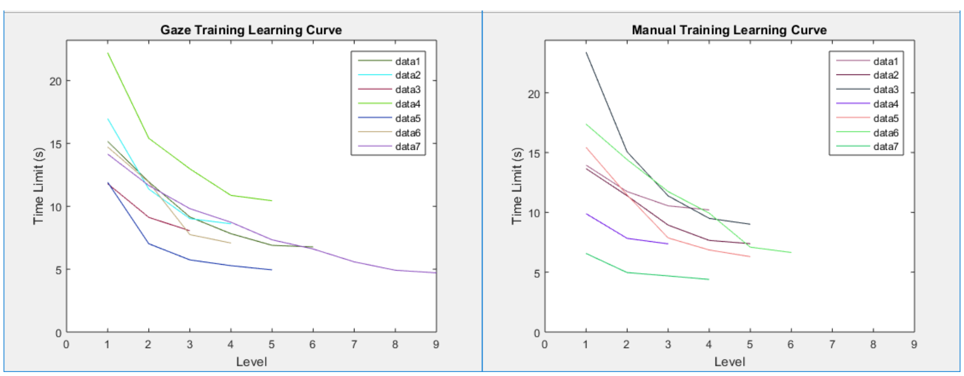
Interestingly, we find that the learning curve is almost identical for both input modalities, as shown in Figure 5.21a, which could be since in many tasks during the gaze-supported learning session, participants did not use the gaze feature at all. On the other hand, we find that the learning curve is in fact different based on the order of the session. Figure 5.21b shows the learning curve by session order. We find that participants, on average, had shorter time limits in the second session compared to the first session, regardless of the input modality. This shows a possibility of carry-over learning effects from one session to the next.

Figures 5.22a and 5.22b show the learning curves for each individual participant per session.



(a) By Input Modality (b) By Session Order

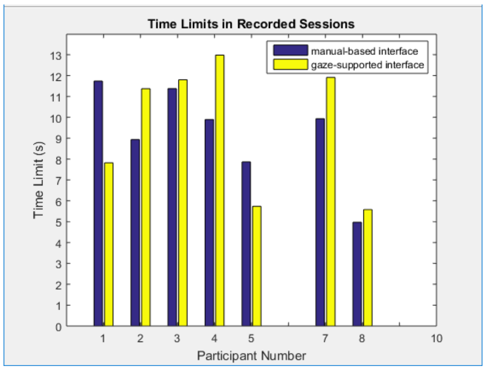
Figure 5.21: Average Training Sessions' Learning Curve per Participant for MZE



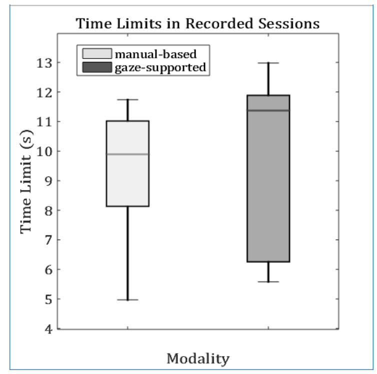
(a) Gaze-supported Training Session (b) Manual-based Training Session

Figure 5.22: Average Training Sessions' Learning Curve for MZE

Recorded Sub-session



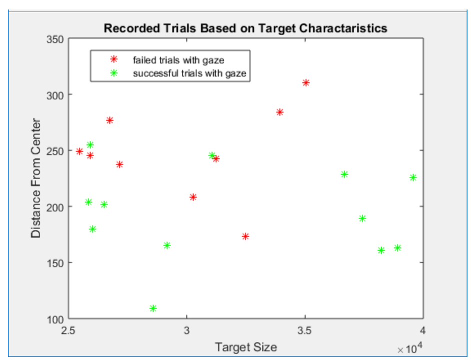
(a) Line Graph



(b) Box Plot

Figure 5.23: Time Limits in Recorded Sessions for MZE

(a) Trials With/Without Gaze Input: Participants tend to use the gaze feature for targets at least 100 pixels away from the centre.



(b) Failed/Successful Trials with Gaze Input

Figure 5.24: Trials Based on Target Size and Distance From Centre for MZE

Like the analysis of One-step Zoom, we look at the time limits for each input modality session. Figure 5.23a shows the time limit per participant per session. Apart from participants 1 and 5, we find that participants had longer time limits for the gaze-supported interface compared to the manual-based interface with much higher deviations. This could be due to the complexity of the gaze-supported interface due to the additional gaze input, which, as described in the participants' post-experiment discussion, takes time to switch to and back to manual. Figure 5.23b shows a box plot for time limits per input modality.

Since participants did not consistently use the gaze feature for all tasks in the gaze-supported recorded sub-session, we base the rest of the analysis on tasks which users used the gaze feature and tasks that didn't. We find that the percentage of times participants decided to use the gaze feature during the gaze-supported session is only 15.91%. Out of these trials, which used gaze, 57.14% of them succeeded and 42.86% of them failed.

Table 5.4 shows the mean time limit for tasks that used gaze and tasks that did not. Both are during the gaze-supported sessions only.

Table 5.4: Mean Time Limit for Gaze-supported Recorded Sessions for MZE

	Used Gaze	**Did Not Use Gaze**
Mean Time Limit (s)	8.06	7.13
Standard Error	0.51	0.23

In addition, Table 5.5 shows in detail the percentage of use of the gaze feature per participant, and how many of these tasks timed out.

Table 5.5: Use of Gaze Feature During Recorded Sessions of MZE

Participant #	**Used Gaze (%)**	**Timed Out (%)**
1	65.22	46.67
2	13.33	0.00
3	17.65	33.33
4	0.00	X
5	0.00	X
7	0.00	X
8	3.85	100.00

We find a highly inconsistent behaviour among participants. For instance, three of them decided not to use the gaze feature at all. One of them used the gaze feature for most of the tasks (65%), another used the gaze feature and had no timeouts and another rarely used the gaze feature and timed out every time.

By inspecting the characteristics of targets which users decided to use the gaze feature for, as shown in Figure 5.24a, we find that participants tended to use the gaze feature for targets that are at least 100 pixels away from the centre. To inspect further, we plot in Figure 5.24b only the tasks in which users used the gaze feature, and highlight the failed trials. We find that most of the failed attempts are the smaller targets, which could be due to the fact that smaller targets need more time for zoom box size adjustment.

5.5.3 Qualitative Evaluation

Due to the small sample size, we report average results in tables and charts. Table 5.6 shows the total TLX scores for both input modalities.

Table 5.6: TLX Scores for Each Input Modality

	Manual-based		**Gaze-supported**	
	Average	**SE**	**Average**	**SE**
Mental Demand	12.50	6.94	18.13	12.92
Physical Demand	17.75	5.62	25.75	5.22
Temporal Demand	54.25	11.47	58.00	11.50
Overall Performance	49.13	9.50	46.88	10.14
Effort	40.38	7.67	51.25	5.92
Frustration Level	14.50	10.19	7.63	3.21
Overall Score	12.57	0.96	13.84	0.55

We find that the overall cognitive load score for the gaze-supported interface is higher than the manual-based interface. By looking at the sources of cognitive load, also graphically illustrated in Figure 5.25, we find that the temporal demand and overall performance contributed most in making the task cognitively demanding for both input modalities. We also notice that the amount of effort participants had to put into the game is higher for the gaze-supported interface, in comparison to the manual-based interface. Contrary to expectations, the amount of physical demand is rated higher for the gaze-supported interface compared to the manual-based interface. This could be since an additional step must be taken to enable eye gaze.

We also find that the manual-based system was more frustrating, on average than the gaze-supported system, with high variability. Considering the sources of frustration, as shown in Figure 5.26, we find that game design-related difficulties caused most of the participants' frustration, such as the

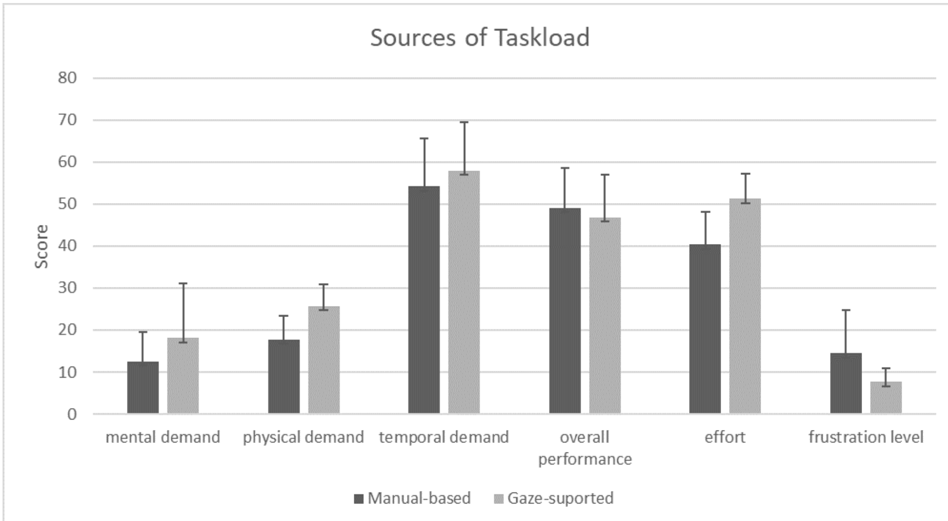


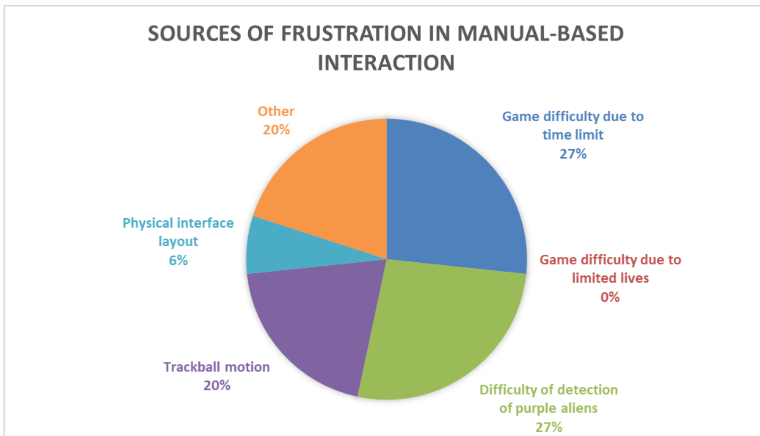
Figure 5.25: Sources of Task Load for MZE

detection of the purple aliens and the time limit. The trackball motion and the physical interface layout also contributed to 26% of the sources of frustration. Participants listed other sources of frustration (20%) as "trying to get the zoom box positioned around the aliens accurately" and "using only the left hand".

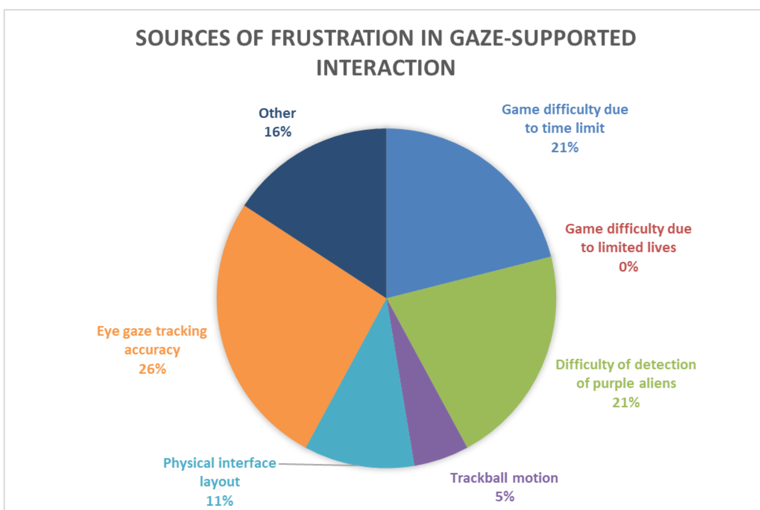
For the gaze-supported system, we find that the lack of accuracy of the eye gaze input forms the highest source of frustration (26%), followed by game design-related sources. Participants listed other reasons as well, such as "toggling between resize and reposition", "trying to get the zoom box positioned around the aliens accurately" and "using only the left hand".

5.5.4 Post-Experiment Discussions

During the discussion session, participants were asked for their general feedback on both systems, and the advantages and disadvantages of both. Given the challenging nature of the game, participants were also asked to elaborate further on their sources of frustration to be able to separate the interaction-related challenges from the gameplay-related challenges. In addition, participants were asked to explain their strategy of using the eye gaze input, since, unlike the One-step Zoom experiment, enabling the eye gaze input feature is optional during the gaze-supported interface session.



(a) Manual-based Interaction



(b) Gaze-supported Interaction

Figure 5.26: The Reported Sources of Frustration During MZE

Advantages of the Eye Gaze-supported Interaction

The only advantages, mentioned by 3 out of 7 participants, is the potential for this eye gaze-supported interface to reduce repetitive stress injuries, as they experienced a lot of repetitiveness with the manual-based interface. In fact, P2 mentioned that he decided to enable eye gaze whenever the game gets frustrating as it provides a little "*physical repetitiveness break*" from the typical manual control. P3 mentioned that *"scrolling is cumbersome: you will have to do two or three scrolls to get somewhere, where you can get there with a button press with eye gaze."* However, this cannot be generalized to all targets, but only to those targets faraway from the centre of the screen. A few other participants mentioned that, had this interaction been applied to larger screens, it would have potentially saved time.

One more advantage of using eye gaze, in this gaming context, is the "fun" factor. Given the added challenge with eye gaze input, participants found it interesting to try out.

Strategies for Using Eye Gaze Input

When asked about the strategies participants used to win the game and when did they decide to use eye gaze as an input, the following were the categories of responses:

- "*I used eye gaze as an input only when the aliens are outside the zoom box.*"
- "*I used eye gaze as an input only when the aliens are at the edge of the screen.*"
- "*I tried to change the size of the zoom box manually and move the box with my eye gaze simultaneously."*
- *"I used eye gaze as an input only because it is fun and novel.*"
- "*I used eye gaze as an input only once or twice before I found out it does not really help.*"

Disadvantages of the Eye Gaze-supported Interaction

We observed that in this user study the disadvantages of using eye gaze input outweighed the advantages. Since this design requires the user to press an extra button to move the zoom box, it adds an extra step, which causes an extra context switch. The mental effort required to perform this switch can mask the benefits of a faster interaction with eye gaze, as the switch itself takes time, as noted by P2.

Once the user has "ideally" used the eye gaze input to perform coarse movements, the user must let go and perform an additional context switch back to the trackball to perform fine movements. This doubles the amount of switching discussed earlier, rendering the eye gaze input unsuitable for time-sensitive tasks.

One of the participants also noticed the little delay at the beginning of enabling the eye gaze input, which is caused due to the initiation of the eye gaze moving average filter, as it requires processing a few fixation points before generating the first filtered fixation point, as explained in Chapter 4.

Like the situation of sonographers, where it is difficult to change their behaviour with the interface they have trained on for years by introducing a new input modality, it is roughly the same case for gamers. One of the participants, who identified as a regular gamer, brought up an important point regarding motor control: *"For gamers, introducing eye gaze for such control takes you out of your comfort zone. I expect gamers will typically resort to switching to the manual controls because that is what they have been trained on for years to win games"* - P5.

5.5.5 Discussion of Results

Although this experiment took longer to finish (MZE took 70 minutes on average per participant, while OZE took 50 minutes on average per participant), results from this Multi-step Zoom user study are not reliable to determine the effectiveness, in terms of time on task and accuracy, of the tested interfaces. A longer, or perhaps a longitudinal, study will be more effective due to the complexity of the gaze-supported multi-step zoom interaction.

Despite the carry-over effect between sessions, the only advantage found in this system is the reduced physical repetitiveness as reported by some participants. However, this is only replaced by another form of physical input that takes up longer time switching to and higher cognitive resources for coordination. Additionally, we find that, in general, participants do not prefer using the gaze feature to win the game, given the added time it takes to activate the gaze feature, in addition to their unfamiliarity with the interaction.

As users gradually decreased their use of the gaze feature during the training sessions, followed by using the gaze feature only 15% of the time on average during the recorded session, we expect that this design will not have promise in terms of an improved interaction. Given these results, we decide not to put emphasis on the MZ interaction in the upcoming clinical user study design as it has not been sufficiently tested in a context-free environment.

5.6 Conclusion

We design a context-free user study to test the proposed gaze-supported interactions (OZ and MZ) independent of external sonography-related effects. The goal of this user study is to collect preliminary results to help us understand the strengths and weaknesses of the proposed interaction design in terms of time on task, accuracy and user behaviour.

Based on these results, we can anticipate the user behaviour for the upcoming clinical user study targeted for sonographers: if the results from this user study turn out to be in favour of eye gaze-supported input, then there is potential for gaze-supported interaction within a more cognitively-demanding context, such as sonography.

We design a multi-modal gaze-supported game, where participants are required to zoom into and destroy targets at particular zoom levels and criteria. We use the same hardware interface layout as ultrasound machines and set similar restrictions regarding left-handed-only interaction. Each participant performs similar set of tasks over two counter-balanced sessions: gaze-supported interaction and manual-based interaction. For each session, participants are required to perform a training task to reach an optimal level of performance before performing the main tasks.

Results from both experiments, OZE and MZE, show a potential in a reduced physical repetitiveness as some manual functions are replaced with a gaze input. However, this reduction comes at the cost of an increased mental demand for both zoom functions and an increased context switch for MZ. In addition, some participants of OZE reported a higher focus on the main task when using the gaze-supported alternative as they manage less manual controls. Quantitatively, both experiments showed high variations in terms of time on task and accuracy. Therefore, we do not have conclusive evidence regarding these metrics.

Results from this context-free user study also revealed some interesting gaze-supported interaction challenges, such as inaccuracies at high zoom levels, an "unintended *Midas touch* effect" and, for MZ, a reduction in the usage of the gaze feature over time. Thus, we expect to encounter these challenges during the context-focused user study. Nevertheless, we expect that testing the interaction with users who are already familiar with the manual interface will eliminate some of the frustration sources faced by the participants of the game user study regarding the unfamiliarity with the base system and cumbersome trackball usage. We also decide to design the context-focused user study with less emphasis on MZ as we were unable to sufficiently evaluate it due to the observed carry-over effects during the training sessions.

Chapter 6

Context-focused User Study: Clinical Experiment

6.1 Goal and Hypotheses

The goal of this study is to test the proposed interaction within the intended context of sonography with end users. In this study, we test the time on task and other eye gaze metrics and relate these quantitative measures to post-experiment discussions regarding the presented system to assess the interaction. We test the following hypotheses:

- **H1:** There is a difference in terms of time on task between the manual-based and the gaze-supported interactions.
- **H2:** There is a difference in terms of cognitive load between the manual-based and the gaze-supported interactions.

6.2 Background on Study Tasks

To design this user study, we first design the tasks to be performed by the participants. The tasks are selected to capture the capabilities of the proposed gaze-supported interaction presented in Chapter 4 within a clinical context. We design this user study with two types of tasks: a realistic ultrasound scan of a healthy volunteer and a number of ultrasound scans of controlled targets.

Replicating a realistic sonography scenario requires selecting a specific ultrasound scan to be performed by the sonographer participants. After discussions with two expert sonographers, an ultrasound scan was selected to be part of this user study based on the following criteria:

- Does not require ill patients, so a healthy volunteer will be suitable for the ultrasound scan.

- The ultrasound scan is common knowledge to all sonographers who perform general ultrasound scans.
- Does not require using a large variety of ultrasound functions, so that the main focus will not diverge from the zoom function.
- The order and number of steps needed to perform this scan does not greatly vary from one sonographer to another.
- Requires capturing a zoomed target.

The selected ultrasound exam is the abdominal routine scans of the common bile duct (CBD) and the common hepatic duct (CHD). This type of exam is performed when a patient requires a general abdominal scan.

Figure 6.1 shows the location of the CBD and CHD and the surrounding structure. The CHD is located before the common hepatic artery and the CBD is located after. In sonography, the goal of this ultrasound exam is typically to measure the size of the CBD and the CHD. The measurement is taken at the largest diameter visible for both ducts. The default size of the CBD and CHD varies based on the gender, age and medical condition of the patient.

The zoom function in this type of exam is important since it is required to perform accurate measurements. The common bile and hepatic ducts are very small structures (3.3 ± 1.1 mm to 6.8 ± 1.1 mm); therefore, measurements should be done at the largest scale possible of the image.

In addition to using the zoom feature, the sonographer frequently changes throughout the exam the depth of field, the gain and the focal zone to obtain the best image. Sometimes the frequency is also changed based on the patient's BMI (Body-Mass Index). The number of focal zones typically used in this type of exam is only one, since the scanned structure is small, horizontal and doesn't stretch vertically across the image. An informal ultrasound exam was observed and video-recorded to identify the exact steps involved in scanning the CBD and CHD, as detailed in Appendix H.

There is no standard or maximum number of images taken of the CBD and CHD to be sent to the physician. The sonographer continues to take images whenever a better view comes up. However, a minimum of 2 images should be taken to show the best measurements.

In addition to testing the interaction through an ultrasound exam performed with a healthy volunteer, we also provide the participant with a medical multi-purpose ultrasound phantom [14]: a specially-designed object used in medical imaging to test the performance of devices in a more

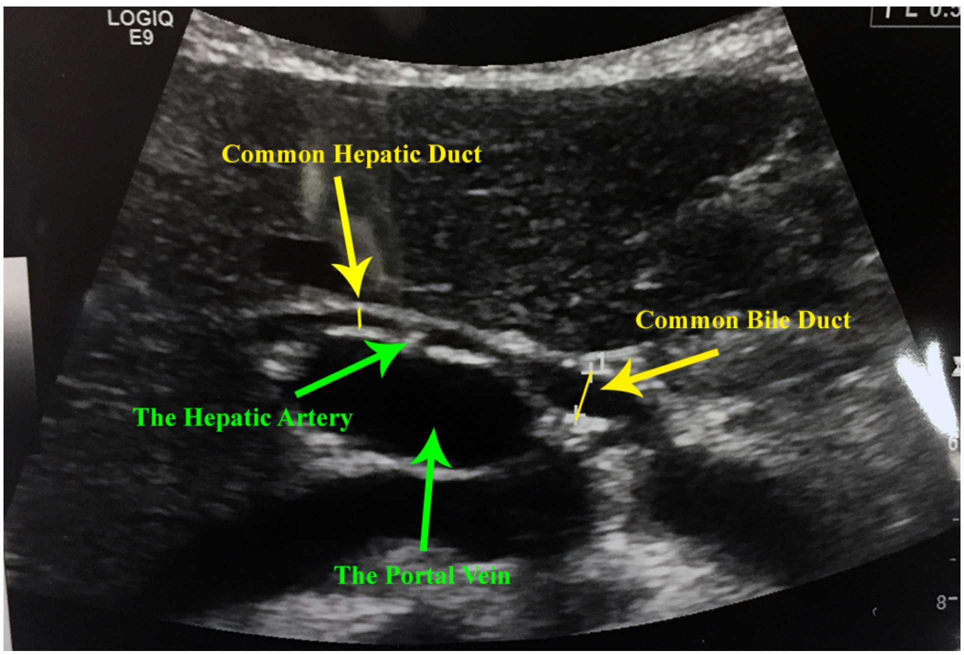


Figure 6.1: An ultrasound image showing the location of the CBD and CHD.

controlled condition instead of using real tissue that is subject to change. Using a phantom also simplifies the task by eliminating the overhead communication between the sonographer and the patient including instructions to change position or hold breathing to obtain better images. Additionally, phantoms enable testing a large pre-defined set of shapes to better understand and accurately measure the performance of the proposed interactions.

We evaluate two different types of phantom target shapes in order to better test the two zoom functions and their different capabilities. Figure 6.4a shows examples of regular target shapes that only require zooming and panning and Figure 6.4b shows examples of irregular target shapes that require zooming, panning and resizing, since a shape's width and height are unequal. In light of the initial results collected from the context-free study, we expect more meaningful results from the OZ interaction, therefore, we provide the participants with more regular target shapes to be acquired with OZ in both training and phantom sessions.

6.3 Experiment

We aim to minimize the differences between the real ultrasound exam setting and the experimental setting, as these differences may cause unwanted learning effects that will mask the real effects of our system. To efficiently test the proposed system, part of the experiment is designed to match a typical diagnostic ultrasound exam setting and match, as close as possible, the typical hardware and software interfaces and room setup of an ultrasound exam as these factors influence the behaviour of the sonographer.

Most of the experiment design decisions were influenced by the implementation and and results of the first iteration of this user-centred design. Details on the structure, procedure, results and lessons learned from the first iteration can be found in Appendix D.

Given that the zoom function is the focus of our test, it will be impractical to run the user study with a full clinical ultrasound exam requiring the sonographer to go through the full steps of a common bile duct exam with measurements. For simplicity, only steps 4 through 10 from Appendix H are applicable in this user study, as requiring the rest of the measurement steps will take away from the focus of this experiment of testing the gaze-supported zoom functions proposed.

As mentioned earlier in Chapter 4, using a specific brand of ultrasound machines in this user study will render the results to be machine-specific, especially that the type of machine available in our labs with an open software interface for data acquisition is not used in typical diagnostic sonography settings. Thus, we use the custom hardware interface we created that only includes the relevant ultrasound functions required for the selected portion of the targeted ultrasound exam: zoom, depth, gain, frequency and focus, in addition to multipurpose buttons around the trackball and the image capture button.

This user study is approved by the Behavioural Research Ethics Board at the University of British Columbia, under UBC CREB number H15-02969.

6.3.1 Setup And Structure

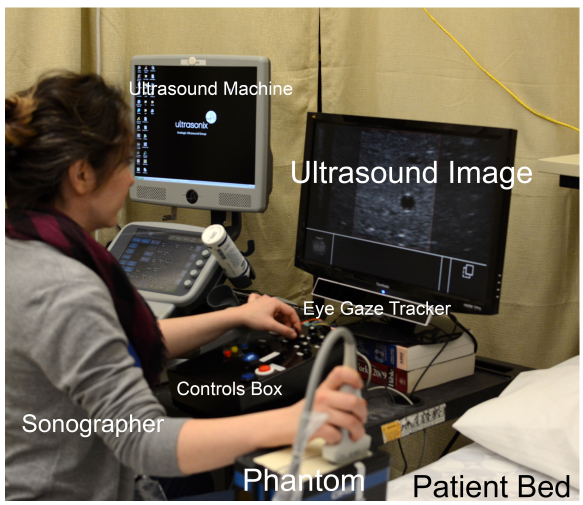


Figure 6.2: The Clinical User Study Room Setup: the setup in the lab closely matches the setup of an ultrasound room in a hospital.

Figure 6.2 shows the setup of the clinical user study. The user study took place at the Robotics and Control Laboratory scan room at the University of British Columbia. The setup consists of the ultrasound machine that streams ultrasound images to an external computer connected to a monitor showing the live ultrasound image. An eye gaze tracker (GP3 Eye Tracker, Gazepoint Research Inc. Vancouver, BC) is mounted at the bottom of the monitor and a custom-made controls box is placed at a lower level in front of the monitor. To the right of the sonographer is the patient bed with the phantom placed over it at a reachable distance to the sonographer. Table 6.1 contains the list of phantom targets scanned by participants. As

Table 6.1: Clinical User Study Phantom Targets and Instructed Techniques of Interaction

Session	**Target #**	**Instructed Technique of Interaction**
Training	1 to 8	One-step Zoom
	9 to 12	Multi-step Zoom
	13 to 16	One-step Zoom
Recorded	1 to 6	One-step Zoom
	7 to 10	Multi-step Zoom

recommended by earlier eye gaze studies [23], eye gaze tracking status is being monitored throughout the whole study for each participant, as eye gaze tracking can be lost during the study for many reasons and require re-setup and re-calibration. The researcher, monitoring the eye gaze of the participant in real time on a separate monitor, is located outside the field of view of the participant to not cause any distractions.

Table 6.2 contains the procedure of the clinical user study. Before running the user study, the room lighting was dimmed to closely resemble the amount of lighting in an ultrasound room. In addition, the settings of the ultrasound image are reset to the default values, to avoid carry-over effects from the settings set by the previous sonographer.

As soon as the sonographer arrives, the sonographer is introduced to the project and is requested to sign a consent form, fill out a demographics form and provided with the participation reward. In addition, the eye gaze tracker is tested and calibrated with the gaze tracker's default 5-point calibration before the user study sessions to make sure there are no eye gaze detection issues, especially in the case of participants wearing highly reflective glasses, as evident in earlier eye gaze tracking studies [23] suggesting the recruitment of 10%-20% more participants than is needed as *"some eye tracking systems may not calibrate well to certain eyes or eyeglasses prescriptions"*. In Addition, *"while most eye trackers claim to work with eye glasses,"* the work in [42] reports, *"we have observed a noticeable deterioration in tracking ability when lenses are extra thick or reflective"*.

Previous work on user-centred ultrasound machine interface design [5] required an analysis of user profiles prior to running the studies, which included collecting information on users, such as the sonographer's experience in ultrasound scanning, i.e. what type of ultrasound scans she performs, experience level and usage patterns of the specific ultrasound function of interest, in this case, the zoom function. We follow a similar approach by collecting this information through a demographics form, found in Appendix J, and a follow-up discussion with each participant sonographer.

Table 6.2: The Clinical User Study Procedure

Session	**Sub-session**	**Task(s)**
Intro	Project Introduction	The researcher introduces the project to the sonographer
	Phantom Exploration	The researcher provides the sonographer with the phantom and requires locating a few random targets.
	Gaze-based Interactions Exploration	The researcher demos the two manual-based zoom interactions (if needed), followed by performing eye gaze calibration of the sonographer, followed by a demo of the gaze-supported zoom interactions.
Training	Perform training tasks using gaze-supported interactions only	The sonographer locates, zooms into and captures the training targets shown in Figure 6.4 and as instructed in Table 6.1.
Phantom	Gaze-supported	The sonographer locates, zooms into and captures the recorded targets shown in Figure 6.4 using only the gaze-supported interactions, as instructed in Table 6.1.
	Manual-based	The sonographer locates, zooms into and captures the recorded targets shown in Figure 6.4 using only the manual-based interactions, as instructed in Table 6.1.
Patient	Locate the CBD	The sonographer explores the patient's abdomen and locates the common bile duct.

Session	Sub-session	Task(s)
	Gaze-supported	The sonographer uses one of the gaze-supported interaction techniques of her choice (or a mix of both techniques) to zoom into and capture 5 consecutive images of the common bile duct of the patient.
	Manual-based	The sonographer uses one of the manual-supported interaction techniques of her choice (or a mix of both techniques) to zoom into and capture 5 consecutive images of the common bile duct of the patient.
Discussion	The researcher discusses the usability of the presented system with the sonographer.	

Instructions on Image Acquisition

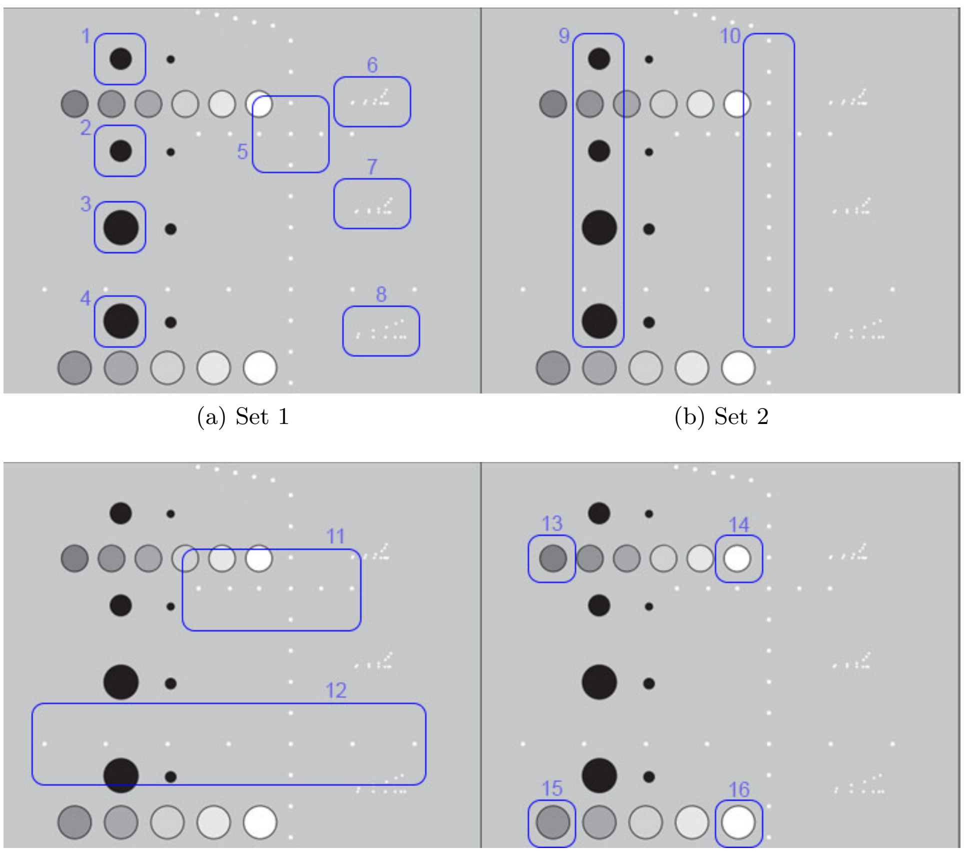
Sonographers are instructed to acquire and capture the zoomed targets filling up most of the image and as centred as possible. They are also instructed to minimize their interaction with ultrasound image settings other than zoom during the recorded phantom and patient sessions.

To help the participants learn the unfamiliar gaze-supported interaction and uncover its capabilities, participants were instructed on the optimal techniques to use the gaze-supported features:

- For the Multi-step Zoom technique, it is best to use the eye gaze feature only when moving the zoom box for long distances across the image.
- Eye gaze is very jittery in small areas, therefore, it is best to use the trackball for fine motions of the zoom box around the target.

Patient Session

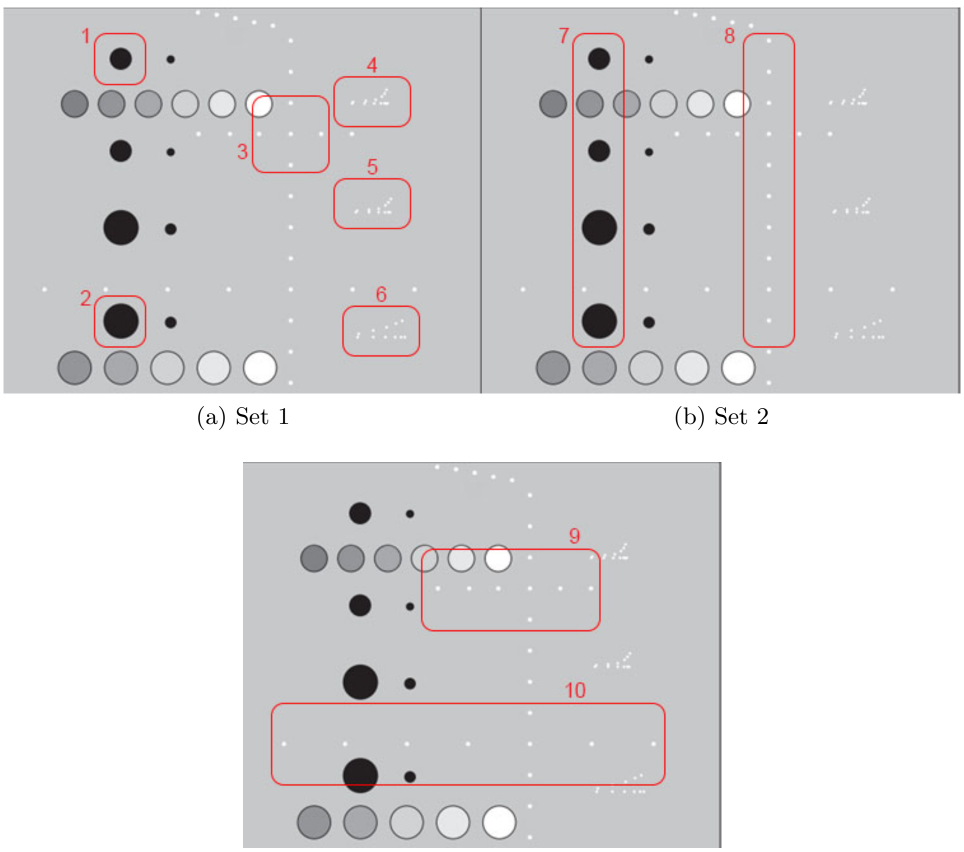
For consistency of the patient target across all trials and participants, the lead researcher volunteered to be scanned by all sonographers. During the



(a) Set 1 (b) Set 2

(c) Set 3 (d) Set 4

Figure 6.3: Phantom Training Targets



(a) Set 1

(b) Set 2

(c) Set 3

Figure 6.4: Phantom Recorded Targets

patient session, volunteers from the RCL lab with a background in eye gaze tracking interaction monitored the eye gaze tracker and conducted the study.

Discussion Session

Sonographers were instructed to provide feedback only at the end of the experiment, during the discussion session, to keep the overall time of the user study under one-hour long.

The following are the list questions discussed with each participant sonographer:

1. What type of ultrasound machine are you familiar with? Did this hardware/software layout and functions closely resemble the ultrasound machines you typically use in your ultrasound scans?
2. What type of zoom do you typically use in your scans? And when? And how frequently?
3. What kind of shapes do you normally need to zoom into?
4. Are they regular shapes with defined centres?
5. How would you describe your own perceived eye gaze behaviour when zooming into targets in an ultrasound image? Do you mainly focus on the target or do you keep peripherally scanning the rest of the image?
6. What advantages and disadvantages did you find using the proposed eye gaze-supported zoom system?

Appendix I includes the full script used by the researcher for the whole user study.

6.3.2 Apparatus

The apparatus used in this user study consists of the same basic tools and setup as the context-free game-based user study presented in Chapter 5. Figure 6.5 shows the hardware setup for this user study. An ultrasound machine, Ultrasonix Touch, transfers the ultrasound images in real time through a TCP/IP connection and communicates the image parameters to the main computer with no observable delays. An eye gaze tracker, Gazepoint GP3, transfers in real time the participant's eye gaze data. The custom-made controls box is operated with an Arduino Mega that is connected to the main computer as well. Figure 4.20 shows the controls box,

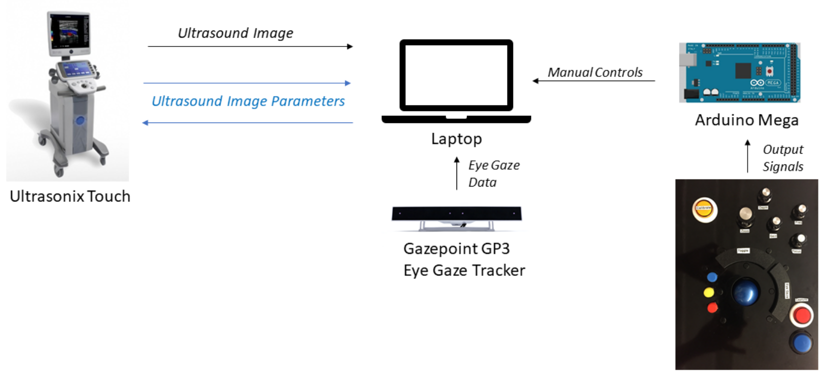


Figure 6.5: The Clinical User Study Hardware Architecture

which has 5 rotary encoders capable of controlling the ultrasound image's depth level, ultrasound beam frequency, b-mode gain level, the horizontal position of one line of focus, and the zoom level. In addition, the trackball is surrounded by the toggle button (used in zoom modes) and the eye gaze button (to enable/disable eye gaze features, as explained in Chapter 4). Finally, the interface also includes a capture button to capture and save the streamed ultrasound image. One missing feature, which was omitted due to technical difficulties during the user study, is the freeze button to freeze the ultrasound image before capturing it. The software interface is fully written in Python using PyQt4 and Pyqtgraph. The communication between the computer and the ultrasound machine is facilitated through a Python wrapper [59] developed at the Robotics and Control Laboratory at the University of British Columbia for Ulterius, a software tool for controlling the Ultrasonix ultrasound systems remotely. Eye gaze data is communicated from the eye gaze tracker through Gazepoint's open Gaze API.

6.4 Analysis Tools

A total of 30 trials per participant were collected and analyzed: 20 phantom trials (10 gaze + 10 manual) + 10 patient trials (5 gaze + 5 manual).

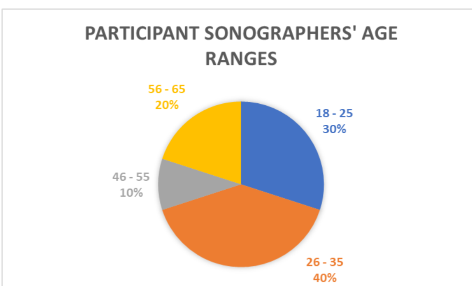
The trials were recorded using Gazepoint Analysis software and manually segmented and transcribed per task. Eye gaze fixation data was produced by Gazepoint Analysis and analyzed by the Eye Movement Data Analysis Toolkit (EMDAT) developed at the University of British Columbia [23].

For each trial, the following dependent variables were recorded and analyzed: time on task and other eye gaze metrics generated by EMDAT: eye movement velocity and fixation rate. In addition, the mean, standard deviation and sum of the following eye gaze metrics were collected: fixation duration and path distance. Descriptions of each of these eye gaze metrics can be found in [21] and [60]. Each task starts from the moment the sonographer has located the target and fixated on it and ends the moment the sonographer captures the target.

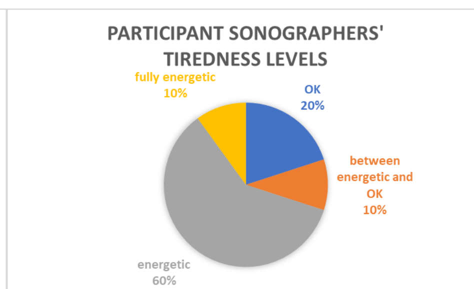
As followed and suggested by previous studies in the field of eye gaze tracking that collected and analyzed similar eye gaze metrics [60], we use Mixed Models for the analysis of variance. Similarly, we apply a Bonferroni correction with m = 4, according to the number of families of dependent variables. This is done to make sure we correct for the family-wise error and multiple-comparisons errors, since there are many dependent variables.

6.5 Results

6.5.1 Demographics



(a) Age Ranges



(b) Tiredness Levels

Figure 6.6: Participant Sonographers' Demographics

A total of ten participants were recruited for the clinical user study. Five of them are professional sonographers and five are student sonographers in their first or second year of their degrees with an experience of operating ultrasound machines and an introductory background of sonography.

Out of the five professional sonographers, 4 are females, of varying levels of experience and age. P1 and P5 have 2-5 years of experience, P4 has 6-10 years of experience, P3 has 20-30 years of experience and P5 has over 30 years of experience. Their ages ranged from 26 to 65 years of age. All recruited professional sonographers routinely perform both general and obstetric/gynaecologic ultrasound exams. All participants, except P5, additionally perform vascular ultrasound exams. P2 and P3 additionally perform MSK, P2 additionally performs cardiac scans, and P4 performs a wider range of ultrasound scans including breast and thyroid imaging.

All student sonographers recruited are female sonographers of the age group 18 to 35 years of age. All student sonographers have an experience in performing both general and cardiac ultrasound exams.

All recruited participants are right-handed with no known abnormal eye or vision conditions. On a tiredness scale of 1 to 5, where 1 is exhausted and 5 is completely rested, nearly all participants selected their tiredness level at 4. None of the participants have used an eye tracker before, except for P3, a professional sonographer, who participated in the previous study conducted in the first iteration of this project. Half of the participants performed the user study wearing glasses, one of the participants was wearing contact lenses, and four had no vision correction. Figure 6.6 visually illustrates the recruited sample.

Consistency of The Presented Prototype With The Manual-based Design of Ultrasound Machines

All recruited participants reported their familiarity with the GE ultrasound machines. In addition, GE machines are used in sonography schools for training students. Therefore, we can eliminate learning effects due to the unfamiliarity of the participants with the basic interaction of pan and zoom and attribute learning to the added gaze feature.

However, subtle exceptions can be made for participants 1 and 5. P1 noted that the default vertical resize function of the trackball is reverse to how it is implemented in a GE machine and it required her some adaptation during the training tasks of the user study to map the direction of the trackball to the reverse mental model of the typical resize function in GE machine. This issue was immediately fixed and the vertical resize direction was consistent with the design of GE machines for the rest of the recruited participants in the study. However, it is important to note that the resize direction mapping can be customized for some models of ultrasound machines, like how the scroll direction in touch-pads can be customized.

On the other hand, P5 required learning the two base zoom functions, as she does not use the zoom feature in her ultrasound scans. Instead, she prefers adjusting the sector size and depth parameters to provide a zoomed and clear image with high definition, which still counts as a form of image magnification.

6.5.2 Observed Gaze-supported Interaction Challenges

With 10 participant sonographers, a total of 10 x 30 = 300 trial samples were collected in total for analysis.

To put the analyzed results in perspective, we identified two classes of challenges related to the completed tasks through the processes of re-observing the recorded trials through Gazepoint Analysis and transcribing them, as summarized in Table 6.3.

Table 6.3: Encountered Challenges During The Clinical User Study Trials

Challenge Classification	**Challenge**
1. System / Interaction Design	1.1. Calibration
	1.2. ROI Lost Before Capture
2. Participant Behaviour	2.1. Multiple Trials
	2.2. Unnatural Forced Gaze Input (Related to MZ)
	2.3. Forced Gaze Input Over Small Areas (Related to MZ)

The System / Interaction Design class describes issues which the user has no control over and are inherent to the way the system is designed and implemented.

- **1.1. Calibration:** refers to a high offset at any area of the screen between the point of gaze detected by the eye gaze tracker and the actual point of gaze of the user.
- **1.2. ROI Lost Before Capture:** we observed that in some cases, the *Midas Touch* problem [32] is not completely avoided, even after using a manual input to activate the eye gaze input. Out of habit, sonographers tend to use some manual inputs and simultaneously look elsewhere outside the screen, either to glance at the manual controls, the probe or the patient. This occasionally happened as the sonographer rotated the zoom knob or held down the gaze input button and

looked away, which caused the system to take a false gaze input and perform zoom into an unintended area, which resulted in losing the region of interest (ROI) and required re-zooming. This resembles the challenge encountered during the context-free user study "*unintended Midas touch effect*", presented in Chapter 5.

Participant Behaviour challenges are those related to the participant not following the instructions or tips provided during the demo session by the researcher for zooming into and capturing targets using the gaze-supported alternatives. These issues could be mitigated by providing the user with enough training to allow a better understanding of the interaction techniques.

- **2.1. Multiple trials:** in some rare cases, the participant confused one target with another for a particular task and had to zoom out and re-zoom to the correct target.
- **2.2. Unnatural Forced Gaze Input:** this occurred in some of the trials of the gaze-supported Multi-step Zoom technique. As participants were provided the option of using the gaze-supported feature to move the zoom box, they sometimes enabled it even when it wasn't needed. For instance, sometimes the participant moved the zoom box manually with the trackball away from the target then used the gaze feature to return it back over the target.
- **2.3. Forced Gaze Input Over Small Areas:** similar to 2.2., this also occurred with the gaze-supported Multi-step Zoom technique. In this case, participants had the zoom box roughly over the correct area, but used the eye gaze input to perform fine movements, which should optimally be adjusted with a trackball instead, as instructed during the demo session of the user study. This attempt to perform slight movements with eye gaze typically offsets the zoom box to an unwanted area, which causes frustration and requires the user to re-perform the placement of the zoom box.

  In our quantitative analysis, we only take into consideration the last correct trial and eliminate all metrics collected from incorrect targets. The researcher always informed the participant of zooming into an incorrect target and requested re-zooming to the correct one.

  We also observed that, especially during the patient session, sonographers always place the ultrasound transducer so that the target structure is in the middle of the acquired image, which in turn does not require much panning of the image after zooming in. This behaviour alone undermines the potential for using gaze-supported zooming, as one of the major advantages, especially for OZ, is the minimization of the required panning as it zooms and pans simultaneously. As a result, sonographers reported not finding a difference in the interaction between gaze-supported and manual-based zooming for patient tasks and some phantom tasks.

6.5.3 Qualitative Results

General Feedback on The Presented System

Potential for Beginner Sonographers One area where the One-step Zoom eye gaze system is found beneficial is in reducing the amount of interaction with the manual controls. Thus, the sonographer can focus more on assessing the acquired image. Especially for novice users, learning the manual controls can take up much of the learner's cognitive attention. P6 and P9, student sonographers, found the potential in using gaze-supported functions in keeping the student sonographer's main attention on the acquired image. As students, one of the frequently occurring scenarios is overly focusing on learning the manual controls, which causes the learner to lose her attention of the probe, causing it to subtly move and lose the area of interest in the acquired ultrasound image.

Added Cognitive Load Associated with MZ With the current capabilities of the gaze-supported MZ, P1 and P2, professional sonographers, did not find the added eye gaze feature an improvement. P1 and P9 reported that the added step to use eye gaze in the Multi-step Zoom function to move the box to the approximate region is a burden rather than a simplification as the sonographer must eventually use the trackball to refine the positioning of the box.

Similarly, student sonographers witnessed some disadvantages with the eye gaze system that could halt its advantages. P7 found it an extra burden that she must remember to switch on the eye gaze feature when needed. P8 noticed that she occasionally moved the zoomed area of interest to an unintentional location as she subconsciously held the gaze-activation button while looking elsewhere either to examine some other feature of the image or to look for the location of a manual input on the manual controls panel.

P2 faced some calibration deterioration issues during the use of the system and stated that the gaze-based system isn't up to an expected level.

"A new interface has to be 10 times better than what we are doing in order to change the habit (of using the manual-based system). It has to be super accurate and very quick." - P2

Negative Reliability One important potential drawback of using eye gaze for fast zooming into images was brought up into attention by P4: it might speed up the interaction to an unwanted level that the sonographer no longer pays attention to the small details in the image. In other words, using a faster zoom function might not necessarily improve the performance or the image content quality.

"When we scan we are basically looking at the whole organ to find minute abnormalities. So, if let's say with the eye gaze it immediately focuses on one thing, you might miss other things." - P4

She also expressed concern for using this function in the long-run, creating an unwanted reliability on the system by the sonographer.

"I think it's helpful but at the same time you can become complacent, because I think that if it is doing the work for me, then I don't have to really focus too much or really search for abnormalities." - P4

Insufficient Amount of Practice with Gaze-supported Input Given the short amount of time sonographers were provided to practice using the gaze-based system, P1, P3 and P5 reported that they found potential in the system, but only if they were provided a sufficient amount of time to practice using it. P4 found it helpful, but very different from the current interaction:

"The eye gaze tracking system is hard to learn but also quite efficient. When you zoom in, the target is just there, so you won't have to keep moving your eyes around." - P4

Given that two functions were improved with eye gaze (One-step Zoom and Multi-step Zoom), participants often mixed up the steps required to use the zoom feature of the two functions, which is an expected behaviour in a short one hour-long user study session.

Phantom Targets are Unrealistic P3 stated that she found difficulty in positioning the zoom box over the phantom task that required framing multiple targets as she had to look in between the two middle targets to position the box. She clarified that in typical ultrasound imaging tasks, there is only one target with a known centre and having to frame multiple targets in one image is unrealistic. After the user study, P5 followed up with an email expressing similar observations:

"It would be unusual for us to zoom in some of the ways outlined in the activity. We would rarely use a long narrow sample box to zoom multiple target areas in a line. Most of the areas we zoom on are singular and use a square box. It is also unusual for us to use a very small sample box and zoom in as close up as we did in the activity. The sample box is rarely less than a quarter the size of the entire sector. We zoom in a bit but always leave information surrounding what we are focused on to provide relational information." - P5

Sonographers' Routine in Ultrasound Scans

Preferred Types of Zoom for Ultrasound Scans Although the zoom function is one of the most frequently-used functions in sonography, based on our field study results presented in Chapter 3, when asked about their ultrasound scans routine outside this user study, the recruited sample of sonographers turns away, in general, from using the zoom function provided with ultrasound machines as the One-step Zoom (known in ultrasound machines as Low-resolution Zoom) degrades the acquired image quality and Multi-step Zoom (known in ultrasound machines as High-resolution Zoom) does not provide an enhanced image quality that competes with other magnification techniques, such as decreasing the depth parameter of the acquired image. P5 clarified further that she prefers adjusting the sector width and the depth parameters to zoom into particular targets located near the surface instead of using the built-in zoom functions of ultrasound machines.

Participants were asked what type of zoom they typically use in their ultrasound scans. Surprisingly, a lot of variations in responses were observed even with the small sample size interviewed, which could be attributed to the different types of exams the sonographers typically perform, their experience with ultrasound machines and their age. Out of ten participants, seven (including all participant student sonographers) typically use Multi-step Zoom instead of One-step Zoom, despite the fact that it takes longer steps to get to, as they are most concerned about the quality of the image. Only one participant prefers the One-step Zoom to the Multi-step Zoom and two participants do not use zoom at all during their ultrasound scans and prefer to use depth adjustments and sector width instead.

"When I was getting my training, we were discouraged from using the zoom function. When you use the zoom function, you cut off the anatomy and you do not see the landmarks." - P2

P1 reported the usage of a mixture of both types of zoom, with a tendency to use the Multi-step Zoom in her ultrasound scans as it provides clearer images than One-step Zoom. P5 stated that she used the zoom function only when she was receiving her overall training on the ultrasound machine as a student, but has never used it afterwards. Similarly, P7, a student sonographer, reported that she typically uses a mixture of both zoom functions, especially when zooming into very small features, such as targets in early pregnancies. For instance, when the Multi-step Zoom cannot zoom any further, she tends to use the One-step Zoom to further magnify the image and accurately place the calipers for doing measurements.

P3's technique is a little different from the other sonographers interviewed: she tends to capture the overall acquired image, then zoom and capture another image of the area of interest. Therefore, the overall contextual image is provided in a separate image for the radiologist to analyze the target within its environment. Supporting the importance of providing a context of the zoomed target, P1 stated *"You want a little bit of context, but you do not want too much surroundings"* as she zooms in to eliminate background noise while leaving some context for the radiologist to locate the target with its environment.

P4 had a completely different preference as she prefers using the One-step Zoom function in her exams instead of modifying the zoomed area parameters as offered by the Multi-step Zoom function. This is due to the fact that she does not substantially zoom into areas in images, but zooms only to differentiate one structure from another.

General Preference of Multi-step Zoom Over One-step Zoom In addition to maintaining the image quality, another reason for using the Multi-step Zoom is that it provides a full visualization of the image before confirming the zoom action. On the contrary, One-step Zoom cuts off the rest of the image as the user gradually zooms in.

"Once you zoom in, slight movements would take you out of the area of interest by moving the probe. And it's very hard to track the area once you have zoomed in already." - P6

In the case of One-step Zoom, sonographers reported that they mainly use it when they generally would like to quickly assess an area without taking further actions. Also, One-step Zoom is used when the image quality and detail is not a priority in the produced image, which is a rare case in sonography. Lastly, it is used to further magnify when Multi-step Zoom reaches its limits, as in the case reported by P7.

Frequency of Use of The Zoom Function The frequency of use is highly patient-dependent as well as target-dependent. All student sonographers reported that zoom is always used in cardiac ultrasound as it is part of their "routine". Other cases also apply for other types of ultrasound exams. For instance, P1 reported using the zoom function when she would like to take an image of the long shape of the endometrium without the distracting surroundings. Other examples are often found in obstetrics and gynecology, such as imaging the fetal heart or particular fetal organs and the ovaries. Sonographers also reported frequently zooming into cists or lesions found in the kidneys, the liver or the breasts. The common bile duct is another example that requires a small level of zoom to enlarge and centre the common bile duct in the image. The zoom function can be used as an intermediate step to another function in the ultrasound machine. For instance, P1 uses the zoom function concurrently with the colour Doppler function when she would like to observe the blood movement in small areas and in small amounts of blood flow. P2 reported that when he used to use the zoom function, it was only to record a zoomed video of the fetal heart to be sent to the radiologist to show that the heart of the fetus is beating.

The zoomed targets can be both regular and irregular. For example, cists and stones are typically circular with a clear centre. P1 stated that some lesions in the liver are highly similar to the phantom targets presented during the study. She also stated that linear shapes are not often found unless they are foreign objects, such as needles or IUDs, or in some cases long blood clots in veins. P4 stated that even shapes that are typically regular in shape can be irregular in some cases. For instance, the liver hemangioma is typically round, but it could be irregular with speckles inside it. Also, the shape of the target is dependent on the posture, position, bowl gas and breathing of the patient.

Participants' Perceived Eye Behaviour In terms of eye behaviour, all interviewed sonographers reported that they tend to scan the whole area then zoom into a particular target of interest. P2 stated that although the sonographer is paying attention to the centre of the image, the sonographer's eyes are actively looking for abnormalities in the peripheral vision. It is also anatomy-dependent: scanning large structures, such as a liver, is different from scanning small structures, such as a thyroid or a lymph node. When scanning large structures, a sonographer's vision is looking everywhere for abnormalities. On the other hand, scanning small structures does not require as much visual attention. In addition, three of the student sonographers reported that they frequently use the context view to help them localize their position in the overall acquired image.

6.5.4 Quantitative Results

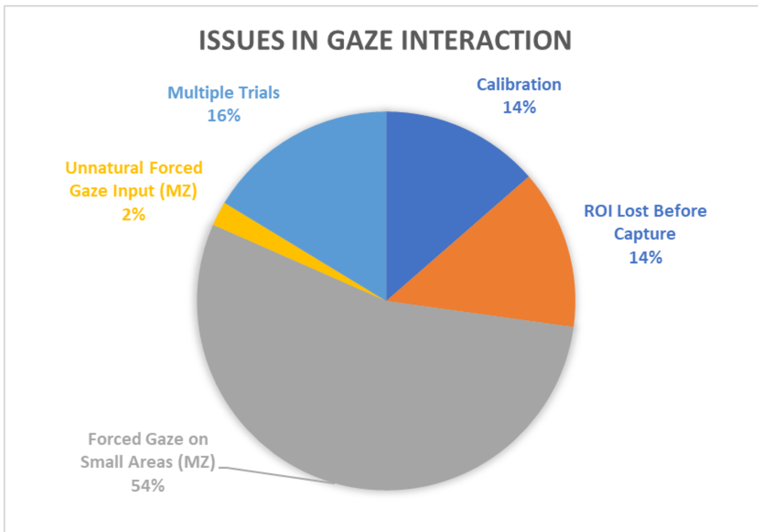


Figure 6.7: The Observed Issues in Gaze Interaction During the Clinical User Study

Out of all gaze-supported tasks, including phantom and patient tasks, 29.6% of the trials exhibited one of the aforementioned observed challenges. Figure 6.7 illustrates a summary.

Gaze-supported Features Usage

We observed an inconsistency in the usage of the Multi-step Zoom technique when the participants were instructed to use the gaze-supported methods. Figure 6.8 summarizes these behaviours: in 45% of the trials during the phantom session and gaze-supported sub-session, participants did not prefer to use the eye gaze feature at all and performed the task with a fully manual-based Multi-step Zoom technique. In addition, in 40% of the trials during this sub-session, participants used the eye gaze input, but incorrectly, or encountered interaction issues. Only 15% of the trials were performed as the gaze-supported Multi-step Zoom interface is intended to be used.

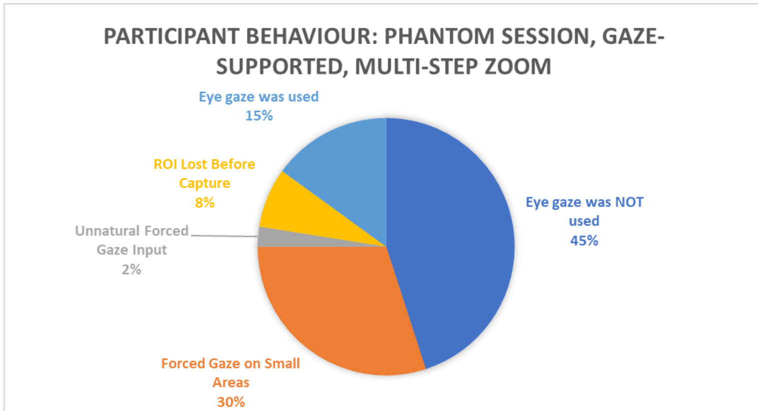
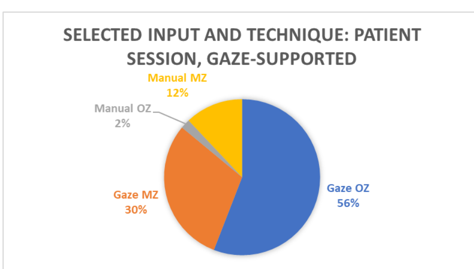


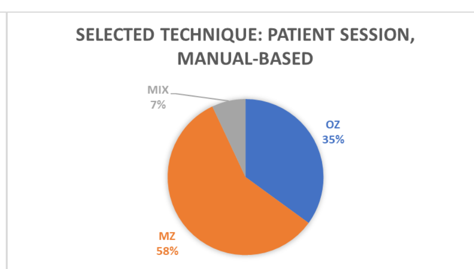
Figure 6.8: Participants' Behaviour During the Phantom Session, Gaze-supported Multi-step Zoom Trials.

Techniques Followed for the Patient Tasks

As participant sonographers were given the choice to select either zoom techniques during the patient tasks, the majority preferred to use the Multi-step Zoom technique when instructed to use the manual-based interface and the One-step Zoom when instructed to use the gaze-supported interface. Figure 6.9 summarizes the participants' choices. Note that, similar to the phantom session, when participants were instructed to use the gaze-supported interface, some preferred to avoid gaze input altogether and replaced it with a manual-based input.



(a) Gaze-supported sub-session



(b) Manual-based sub-session

Figure 6.9: Participants' Choice of Input Method and Technique During the Patient Session

6.5.5 Results from the Mixed Models Analysis of Variance

A 2 (interaction: Gaze-supported, Manual-based) x 3 (Zoom Technique: OZ, MZ, Mixed) x 3 (Target Type: Regular Phantom, Irregular Phantom, Patient) Mixed Models Within-subject ANOVA was performed on the collected data, with two discarded trials out of 300, due to missing eye gaze data.

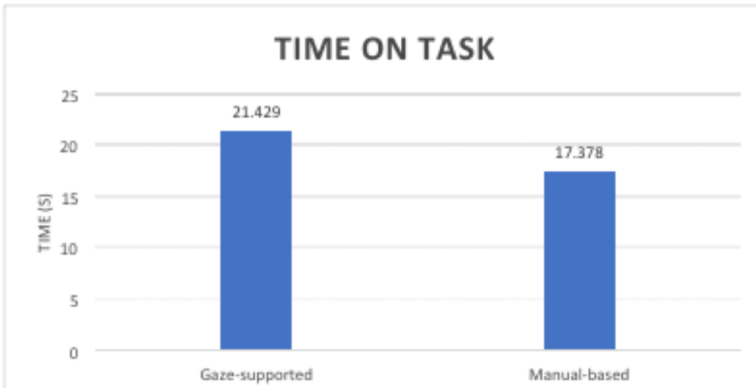
By regular phantom targets, we refer to targets that have a uniform shape in the phantom, such as the training phantom targets 1 to 8 and 13 to 16. By irregular phantom targets, we refer to multiple uniform shapes in a row (or a column), such as the training phantom targets 9 to 12.

Table K.1, found in Appendix K, includes a summary of post-processing applied on the collected data, including transformations and number of trimmed outliers to correct violated assumptions and reports on the violated assumptions that persisted even after data post-processing. Outliers are trimmed if they are above or below 2.5 of the standard deviation.

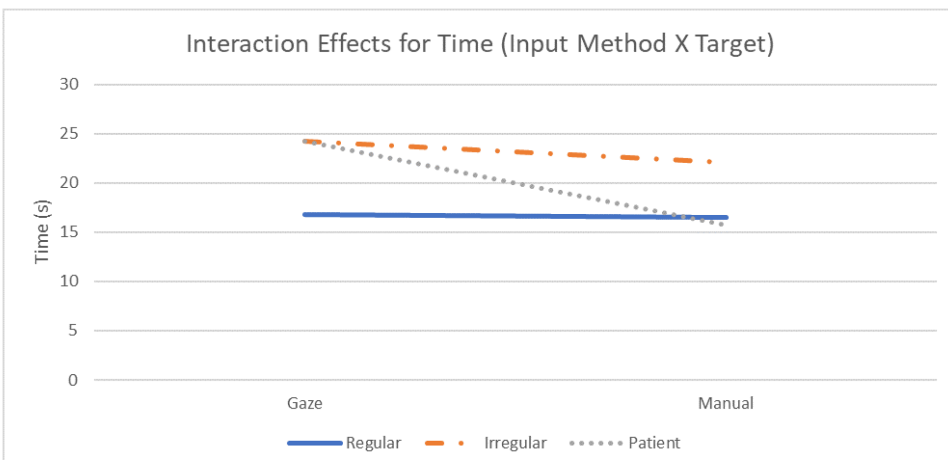
Effects of Input Method

Time on Task. There was a main effect of Input Method on Time on Task. $F(1, 278) = 22.91$, $p < 0.001$. Pairwise comparisons showed that users spent significantly more time on task using the gaze-supported input method (Gaze-supported $M = 21.43$ sec. $SD = 0.22$, Manual-based $M=$ 17.38 sec, $SD = 0.20$), as shown in Figure 6.10a. This result aligns with the fact that some participants did not find the provided user study time of one hour in one day enough to train on using the gaze-supported tasks. Therefore, we hypothesize that much of the time difference between the two input methods is attributed to the learning effect.

There was an interaction effect between Input Method and Target on Time on Task. $F(2, 275) = 9.91$, $p < 0.001$. Bonferroni post-hoc tests examining the interaction effects found that there is no significant difference in terms of performance between the two input methods when the target is the phantom. The source of variation is largely due to the patient CBD target, with a difference of $\Delta = 8.54$ seconds, $p < 0.001$. This means that when sonographers are faced with a patient task, their performance using the gaze-supported interaction drops by 54.5%. This drop in performance could be attributed to a number of factors, including the context switch required to communicate with the patient: in the case of gaze-supported input, the sonographer has to be aware of her gaze at all times, which makes the context switch between the ultrasound image and the patient



(a) Main Effect By Input Method



(b) Interaction Effect Between Input Method and Target

Figure 6.10: Time on Task Statistics

more demanding. It also could be attributed to the fact that 30% of the selected techniques used during the patient gaze-supported sub-session is the gaze-supported Multi-step Zoom, which takes significantly longer to use compared to the gaze-supported One-step Zoom, as will be discussed in later results. By referring to Table 6.4, we further observe that gaze-supported interaction performs the best with regular phantom targets, which are often captured with One-step Zoom. A significant difference, p <0.001, between scanning regular phantom targets and patient targets when using the gaze-supported input, and an identical mean of time on task performance for irregular phantom targets and patient targets support this finding.

On the other hand, Table 6.4 also shows that the manual-based interactions perform the best with patient targets with a significant difference, p = 0.008, from the irregular phantom targets, which could be due to the familiarity of the participant with both, the input method and the target. The second-best target scanned with a manual-based input is the regular phantom targets, with a significant difference from irregular phantom targets, p = 0.045.

This low performance with irregular phantom targets could be due to the fact that the shape of the irregular phantom targets is unusual for sonography tasks, as reported earlier by P3 and P5.

Table 6.4: Mean Time on Task Based on Input Method and Target

	Input Method	**Gaze-supported**	**Manual-based**
Target	Phantom Regular	16.75 s	16.44 s
	Phantom Irregular	24.21 s	22.01 s
	Patient	24.21 s	15.67 s

Eye Movement Velocity. No main effect was found of input method on eye movement velocity. However, there was an interaction effect between Input Method and Target on Eye Movement Velocity. $F(2, 275) = 5.01$, p = 0.028. Bonferroni post hoc tests examining the interaction effects found that users' eye velocities are significantly different across input methods only when users were scanning regular phantom targets (p = 0.001). The interaction is illustrated in Figure 6.11.

Fixation Rate. No main effect was found of input method on fixation rate.

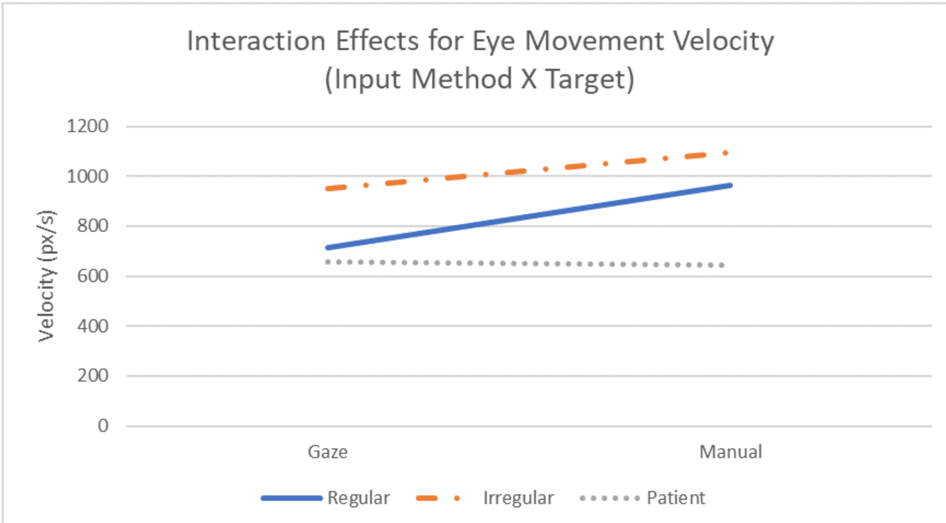


Figure 6.11: The Interaction Effect Between Input Method X Target on Eye Movement Velocity: with higher zoom levels, using the gaze-supported interaction slows down the eye movement velocity.

Mean Fixation Duration. There was a main effect of Input Method on Mean Fixation Duration. $F(2, 274) = 17.28$, p <0.001. Pairwise comparisons showed users' mean fixation durations are significantly higher using gaze input methods (M = 483.62 ms, SD = 1.26) compared to manual methods (M = 427.27 ms, SD = 1.28), as shown in Figure 6.12. This result aligns with the fact that some participants found the gaze-supported zoom methods more cognitively demanding, as some research in eye tracking suggests that longer fixation durations is an indication to higher cognitive load due to the allocation of the cognitive capacity to information processing [49] [63].

Mean Path Distance. No main effect was found of input method on mean path distance.

Effects of Technique

Time on Task. There was a main effect of Technique on Time on Task. $F(2, 281) = 5.24$, $p = 0.024$. Pairwise comparisons showed that users spent significantly less time using the OZ technique compared to the MZ technique

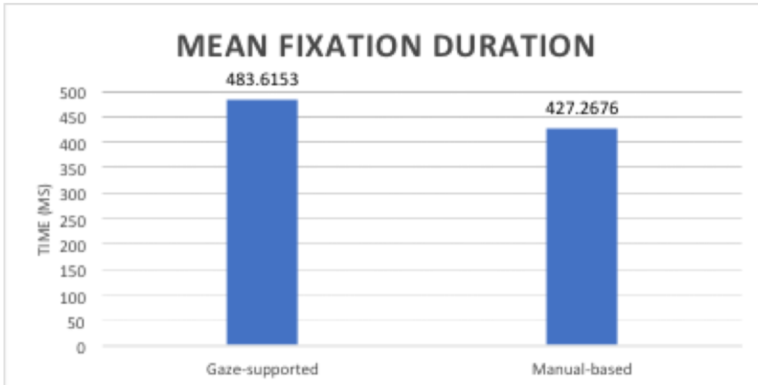


Figure 6.12: Main Effect on Mean Fixation Duration By Input Method

(OZ M = 17.18 sec., SD = 0.20, MZ M= 21.09 sec., SD = 0.19). This is mainly due to the fact that MZ requires more steps than OZ to achieve the required magnified image, which consequently takes longer to achieve.

Eye Movement Velocity. There was a main effect of Technique on Eye Movement Velocity. $F(2, 282) = 7.13$, $p = 0.004$. Pairwise comparisons showed users' eye velocities are significantly faster using MZ (M = 914 px/s, SD = 346.73) compared to OZ (M = 724 px/s, SD = 373.42). This is likely to be since the sonographer rapidly moves her eye gaze around the edges of the zoom box to ensure the correct placement around the area of interest.

Fixation Rate. No main effect was found of technique on fixation rate.

Mean Fixation Duration. No main effect was found of technique on mean fixation duration.

Mean Path Distance. There was a main effect of Technique on Mean Path Distance. $F(2, 270.68) = 10.62$, p <0.001. Pairwise comparisons showed users' fixations mean path distances are significantly longer using MZ (M = 556.67 px, SD = 174.95) compared to OZ (M = 356.60 px, SD = 183.66), p <0.001. This is due to the fact that participants are instructed to use MZ with irregular targets, which cover up more space within the ultrasound image compared to regular targets that are zoomed with OZ.

Effects of Target

Time on Task. There was a main effect by Target on Time on Task. F(2,278)=5.70, p = 0.016. Pairwise comparisons showed that users spent significantly less time to capture the regular phantom targets compared to the irregular phantom targets (Phantom Regular M = 6.56 sec., SD = 0.20, Phantom Irregular M = 23.12 sec., SD = 0.18). Similarly, this is due to the fact that most of the regular targets were zoomed in using OZ, which takes significantly less time to use than MZ.

In addition, differences were found when using the manual input method depending on the type of target scanned. Users took significantly longer scanning irregular phantom targets in comparison to regular phantom targets (p = 0.012) and took significantly longer scanning irregular phantom targets compared to patient targets (p = 0.002). Figure 6.10b illustrates these interaction effects.

Eye Movement Velocity. There was main effect on Eye Movement Velocity was by Target. F(2, 277) = 15.54, p <0.001. Pairwise comparisons showed that users' eye velocities are significantly slower when scanning patient targets (M = 650 px/s, SD = 320.02) compared to regular phantom targets (M = 863 px/s, SD = 401.87). Similarly, velocities are lower scanning patient targets compared to irregular phantom targets (M = 1022 px/s, SD = 315.92). This decrease in speed in eye movement could be due to the sonographers' analysis and close examination of the CBD. On the other hand, phantom targets do not require much examination as they are not realistic patient targets that need analysis.

Fixation Rate. There was a main effect of Target on Fixation Rate. F(2, 421767) = 6.22, p = 0.008. Pairwise comparisons showed users' fixation rates are significantly higher scanning irregular phantom targets (M = 1.90 fix/s, SD = 0.32) compared to scanning patient targets (M = 1.60 fix/s, SD = 0.37).

Mean Fixation Duration. No main effect was found of target on mean fixation duration.

Mean Path Distance. There was a main effect of Target on Mean Path Distance. F(2, 278)= 7.15, p = 0.004. Pairwise comparisons showed users' fixations mean path distances are significantly longer scanning phantom regular (M = 424.58 px, SD = 216.83) and phantom irregular (M = 579.54 px,

$SD = 155.04$) compared to patient targets ($M = 395.18$ px, $SD = 183.21$), with p values of <0.001 and equal to 0.002 respectively. Again, this is possibly due to the familiarity of the sonographer and the nature of repetitiveness of the patient tasks compared to the phantom tasks.

6.5.6 Suggested Improvements for Other Ultrasound Machine Functions

Despite the varied results received from participants, sonographers found potential in the technology to improve other functions in the ultrasound machine after testing the capabilities of the eye tracking technology integrated with the machine. P2 suggested a use for the eye gaze feature to help identify the anatomy the sonographer is looking at and suggest annotations with the help of some image recognition feature. Another suggestion provided by P2 is to use the eye gaze- supported features to help inspect larger ultrasound images, such as panoramic images. P5 suggested using eye gaze with ultrasound machines as a teaching aid for teaching student sonographers what areas they should be inspecting and what tissues should be assessed. Similarly, this eye gaze can be recorded in practice and provided to radiologists to highlight the areas the sonographers were inspecting during the scan. Additionally, the same approach for setting the position of the zoom window of the Multi-step Zoom function can be adopted for moving the Doppler window around the screen automatically based on eye gaze.

6.5.7 Discussion of Results

Based on the quantitative results presented, we reject the null hypothesis regarding time on task (**H1**). Similarly, provided the reported qualitative results, along with the longer fixation times, we reject the second null hypothesis regarding cognitive load (**H2**). Through our user study, we found that gaze-supported interaction takes significantly (23.3%) longer than manual-based interaction. Similarly, participants reported a higher cognitive load associated with the gaze-supported solutions, especially for the Multi-step Zoom function. Mean fixation durations are higher by 13% using the gaze-supported input compared to the manual-based input. However, given the novelty of the interaction, we acknowledge that these results are not conclusive, as the interaction has been tested for one day, under one hour per participant.

Using a gaze-supported zooming approach in diagnostic sonography has its potential only in niche areas. As noted by student sonographers, this could be helpful to alleviate the distractions caused by learning the manual ultrasound controls for beginners and allow them to focus more on the presented ultrasound image. For expert sonographers, this benefit makes no difference as they are very skilled at operating the ultrasound machine's manual controls.

In fact, our results regarding longer time on task performance and higher mental demands, in addition to the qualitative feedback collected from the participant sonographers, suggest that complications arising from using the gaze-supported interface in the long run will hinder the benefits intended by the design of the gaze-supported solution in terms of lowering the repetitive strain injuries for ultrasound machine users by dividing the input between the motor control channel and the visual channel.

Limitation

Evaluating ultrasound machine interfaces is of particular challenge, as found in earlier studies [4], as there are many factors influencing the interaction, including, but not limited to: the user's level of experience, the user's attitude to the product, the clinical application, the clinical work flow and the type of ultrasound system tested. The challenge is even manifested as we have limited access to users. Therefore, we acknowledge that there could be some external factors influencing our results. The most important factor is the tendency of the recruited participant sonographers to use MZ over OZ in their daily routine scans (high-resolution zoom is preferred to low-resolution zoom) or other mechanisms to magnify ultrasound images.

6.6 Conclusion

We present a context-focused clinical user study designed for sonographers to assess the proposed gaze-supported zooming interaction quantitatively and qualitatively in terms of time on task and added cognitive load. We base the user study design on the routine CBD scan performed by sonographers provided its simplicity and familiarity by all expert and student sonographers. We include phantom targets in the structure of the user study as well to examine the behaviour of sonographers and the performance of the system with a variety of controlled shapes.

The user study takes place at our lab with an environment setting closely matching to the ultrasound room at a hospital. The user study is structured so that each participant receives an equal amount of training using the gaze-supported zoom functions before performing the user study tasks. The user study tasks involve an equal number of tasks to be performed with the manual-based and gaze-supported zoom interaction. They also involve tasks to be performed by One-step Zoom and Multi-step Zoom. The user study tasks are followed by a discussion with the recruited participant sonographer to qualitatively evaluate the presented gaze-supported system and its potential compared to the traditional manual-based system.

A total of five expert sonographers and five student sonographers participated in the user study with varying levels of experience. Through our results analysis, we observe five frequent gaze-supported interactions related challenges, which occurred during 29.6% of the gaze-supported tasks.

Sonographers found potential in the tested gaze-supported interaction for training student sonographers, as it alleviates the need to focus on the physical input layout and allows for higher focus on the ultrasound image. However, they also report an increased cognitive demand, especially when using Multi-step Zoom, due to the novelty of the interaction and the need for the users to "be aware of where they are looking" at all times when the gaze is being actively used as an input.

In alignment to the results obtained from the context-free game-based user study, we find that gaze-supported OZ is significantly faster than gaze-supported MZ, used more often by participant sonographers during the gaze-supported patient session and shows no significant difference in terms of mean fixation duration. Therefore, out of the two techniques, we find that gaze-supported interaction performs better when it is implicitly integrated as a control input.

We observe that there could be external factors influencing our results. The most important factor is the tendency of the recruited participant sonographers to use the Multi-step Zoom in their daily routine scans or to avoid using the zoom function altogether and use alternative mechanisms to magnify ultrasound targets. Another factor is the insufficient exposure to the gaze-supported system and the lack of training.

As for manual-based interaction outside the presented user study, sonographers tend to use Multi-step Zoom over One-step Zoom as it does not degrade the resolution of the zoomed image. Had ultrasound machines provided high resolution images with the One-step Zoom approach, this issue would not be a concern anymore for the sonographer's preference of the Multi-step Zoom interaction over the One-step Zoom. However, some sonographers might still prefer MZ over OZ, as it provides higher visualization, as discussed earlier.

Chapter 7

Conclusions And Recommendations

7.1 Conclusions

In this thesis, we follow a user-centred design approach to investigate, design and evaluate two multi-modal gaze-supported zoom interactions, Multi-step Zoom and One-step Zoom, for zooming into the acquired images in ultrasound machines. We define the zoom functions in ultrasound machines, the High-resolution zoom and Low-resolution zoom, as a subset of a larger group of functions concerned with image magnification and analyze the user interaction to create an informed gaze-supported interface design.

We present a complete state-based analysis of zoom functions, OZ and MZ, in ultrasound machines and integrate gaze tracking capabilities. We test our presented gaze-supported zoom interactions through a series of user studies. Results from the context-free game-based user study helped us identify the potential improvements and challenges of using the investigated gaze-supported interaction techniques. We observed that both gaze-supported techniques required lower physical demand at the cost of introducing a higher mental demand compared to the manual-based techniques. In addition, other challenges were observed, such as inaccuracies at high levels of zoom and an "unintended *Midas touch* effect" in both gaze-supported techniques and a higher context switch in the gaze-supported MZ.

A total of five expert sonographers and five student sonographers participated in the context-focused user study. We observe five frequent interactions-related difficulties, which occurred during 29.6% of the gaze-supported trials. In our results, we find that gaze-supported interaction requires significantly higher time on task and longer fixation duration compared to manual-based interaction. This indicates that using the presented gaze-supported alternatives is slow and cognitively demanding for the selected sonography tasks. Sonographers report an increased cognitive demand due to the novelty of the interaction and the need for them to "be aware of where they are looking" at all times when the gaze is being actively used as an input.

However, participants in both user studies reported that they experienced a higher focus on the main tasks when using the gaze-supported OZ technique compared to the manual-based alternative. This is because their attention was not occupied with managing the manual controls. In addition, sonographers found potential in the presented gaze-supported interaction for training student sonographers, as it alleviates the need to focus on the physical input layout and allows for higher focus on performing their main task of analyzing the ultrasound image.

To compare the two zoom-supported interaction, One-step Zoom shows higher potential than Multi-step Zoom, as it is proven to perform faster, depending on the targets. On the other hand, gaze-supported Multi-step Zoom added to the complexity of the task that is attributed to the repetitive context switch required to activate and deactivate the eye gaze input to move the zoom box. In other words, we find that Multi-step Zoom violates the rule of using implicit gaze input, as detailed in Chapter 4, as the user ends up actively controlling the position of the zoom box on the screen. Conversely, One-step Zoom does not place that much impact on actively controlling an object on screen, but implicitly uses the location of the user's eye gaze to zoom into an area of interest.

Since both user studies were performed only for one hour per participant, higher times on tasks could be attributed to the fact that the participants did not have enough exposure to and experience using the presented gaze-supported zooming interface. Further evaluation is recommended to test the interaction independent of this learning effect. We recommend a longitudinal study performed with novice sonographers extending to multiple days.

Given these results, we find that there is no substantial evidence that gaze-supported zooming is beneficial in ultrasound machines in terms of improving speed and mental workload. However, the observed potential of the gaze-supported OZ technique in terms of higher focus on tasks and reduced physical strain could be further investigated in followup studies.

7.2 Recommendations

To take this work further, we recommend another iteration with a longitudinal study performed with student sonographers in their first year of training to test the effect of the interaction over a longer period, as participants in both iterations were exposed to the gaze-supported system for less than an hour.

For the next iterations, we recommend extending the set of eye gaze features we looked at to bring more insight and understanding with regard to users' cognitive load. For instance, we recommend examining pupil dilation, since it is another known gaze feature that changes with mental workload [7]. In addition, we recommend investigating Area of Interest (AOI) features, which will allow checking for differences in how users process specific regions of the interface, e.g., target set shapes.

In addition, we recommend performing formal evaluation of ergonomics of the new interaction, as it is of main concern in the field of sonography, as discussed earlier. Ultrasound machine interface ergonomics evaluation techniques are discussed in [3] and [64], including motion analysis, superficial electromyography, digital human modelling and observational studies with camera recording.

In terms of system design, one possible modification to the fast zoom function would be to reduce the sensitivity gradually (or to stop it altogether) once the image reaches a certain amount of zoom, as the error gets larger as the zoom level is higher.

In addition, other areas can be explored in the ultrasound machine to investigate the integration of eye gaze tracking. Another object dragging task, although 1-dimensional, that will benefit from eye gaze support, is the automatically setting the focus level to the area where the sonographer is interested in (i.e. where the point of gaze is located). However, automatically changing the focus levels based on the eye gaze will lead to the *Midas touch* problem: the sonographer might simply be inspecting a particular depth of the image, not intending to set the focus levels to it. Therefore, similar to the discussed object-dragging task in Chapter 4, a muscle group has to be engaged while setting the focus levels. This could be implemented simply with depressing a button to set the focus levels. The task is still 1-dimensional, but the manual input device required is no longer a 1-dimensional input, but a binary input supported with the implicit input of the eye gaze.

In Zhai's MAGIC work [68], hotspots were identified in the user interface. If the user's eye gaze is in the vicinity of a hotspot, it is automatically drawn into it, which makes recognition easier for areas where the user might be looking at. For future work, the same approach can be applied by performing image processing on the ultrasound images to recognize the "hotspots" or areas that the sonographer is likely to perform ultrasound functions on (magnify, measure, etc.), which could improve the performance of eye tracking-supported ultrasound functions.

Other areas where eye gaze interaction can be of help is in the development of gaze-based menu selection for ultrasound image parameters for hands-free cases, such as for ultrasound-guided interventional procedures, in case of absence of an ultrasound operator to assist the radiologist in performing the procedure. Earlier work in eye gaze tracking explored gaze-based menu selection interaction, such as that presented in [36]. A popular approach for gaze-based selection techniques is gaze gestures, as presented in [16], which shows how it can be integrated into standard desktop applications. Work such as presented in [52], explores automatic cropping of images based on eye gaze fixations. This work is completely gaze-based and does not use a secondary input modality to support the interaction. This work can be beneficial to develop automatic zoom methods for hands-free ultrasound machine interaction, such as when used in ultrasound-guided procedures and surgeries. Given the frequent context switch that sonographers undergo during an ultrasound scan, it might also be worth exploring eye gaze context switch facilitation techniques, such as that presented in [38], to aid the sonographer in switching from the ultrasound image and back to the same area that was being examined.

7.3 Contributions

Our contributions in the work presented include the following:

1. We presented results from a field study that includes observations of routine diagnostic ultrasound exams at two hospitals, interviews with sonographers, and a survey to members of the British Columbia Ultrasonographers' Society. Through the field study, we understand the context of sonography, the ultrasound machine functions used during ultrasound exams and the challenges faced by sonographers, including the encountered WRMSDs. Thus, we identified the potential advantages and the starting point to using a gaze-supported multi-modal ultrasound machine interface and the potential risks associated with it.

2. We presented an analysis of the magnification-related functions and the tasks associated with them during ultrasound exams through a combination of results obtained through the field study and the user studies. We find that magnification is always needed in ultrasound exams. However, the approach to achieving magnification can vary. Our field studies show that zoom is one of the most frequently-used functions in ultrasound exams. Our interviews with sonographers during

the context-focused user study show that sonographers also use different approaches to magnify targets, including controlling the image depth and sector size. In addition, the type of zoom function used, High-resolution or Low-resolution Zoom, highly depends on the application, the target, and the preferences of the sonographer.

3. We presented a state representation of the manual-based zoom functions available in the class of ultrasound machines used in routine diagnostic sonography to better understand and visualize the interaction. Our state representation shows that the zoom functions analyzed implement all or a combination of three states: Full-scale, Pre-zoom and Zoom.

4. We presented a modified state representation of the zoom functions, which integrates the gaze input, implicitly and explicitly, to create a multi-modal gaze-supported zoom interaction for both High-resolution and Low-resolution Zoom (or what we refer to as Multi-step Zoom and One-step Zoom). This state representation was achieved through several iterations. We also present the implementation of the proposed gaze-supported interaction.

5. We analyzed the interfaces of the recent premium to high-end ultrasound machines used in routine diagnostic ultrasound and observed during the field study and built a custom hardware controls interface that resembles the same design patterns analyzed and contains only the main functions needed for evaluating the gaze-supported zoom interactions.

6. We presented results of a context-free game-based user study designed to evaluate the interaction with the proposed gaze-supported interface in isolation to external effects related to sonography. Results from this user study showed a potential for a reduced physical interaction and an improved focus on the main tasks. The results also anticipated risks related to an unimproved time on task, accuracy and cognitive load and identified other observed behaviours and challenges related to gaze-supported interaction.

7. We presented results from a context-focused clinical-based user study performed with sonographers, which showed an increased time on task and cognitive load and identified other areas of improvements where this type of interaction has potential, such as to aide student sonographers in learning the machine controls and focusing on the main tasks,

given the reduced amount of manual inputs with the gaze-supported interface.

Bibliography

[1] Chadia Abras, Diane Maloney-Krichmar, and Jenny Preece. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4):445–456, 2004.

[2] Nicholas Adams, Mark Witkowski, and Robert Spence. The inspection of very large images by eye-gaze control. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '08, pages 111–118, New York, NY, USA, 2008. ACM.

[3] G. Andreoni, M. Mazzola, S. Matteoli, S. DOnofrio, and L. Forzoni. Ultrasound system typologies, user interfaces and probes design: A review. *Procedia Manufacturing*, 3:112 – 119, 2015. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015.

[4] G Andreoni, M Mazzola, S Matteoli, S DOnofrio, and L Forzoni. Ultrasound system typologies, user interfaces and probes design: A review. *Procedia Manufacturing*, 3:112–119, 2015.

[5] Arlene F Aucella, Thomas Kirkham, Susan Barnhart, Lawrence Murphy, and Kris LaConte. Improving ultrasound systems by user-centered design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 38, pages 705–709. SAGE Publications Sage CA: Los Angeles, CA, 1994.

[6] Patrick Baudisch, Nathaniel Good, Victoria Bellotti, and Pamela Schraedley. Keeping things in context: A comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 259–266, New York, NY, USA, 2002. ACM.

[7] Jackson Beatty and Brennis Lucero-Wagoner. The pupillary system. *Handbook of psychophysiology*, 2:142–162, 2000.

[8] Aga Bojko. *Eye tracking the user experience: A practical guide to research*. Rosenfeld Media, 2013.

[9] Andreas Bulling and Hans Gellersen. Toward mobile eye-based human-computer interaction. *IEEE Pervasive Computing*, 9(4):8–12, 2010.

[10] William Buxton. Lexical and pragmatic considerations of input structures. *ACM SIGGRAPH Computer Graphics*, 17(1):31–37, 1983.

[11] William Buxton. A three-state model of graphical input. In *Human-computer interaction-INTERACT*, volume 90, pages 449–456, 1990.

[12] Ishan Chatterjee, Robert Xiao, and Chris Harrison. Gaze+ gesture: Expressive, precise and targeted free-space interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pages 131–138. ACM, 2015.

[13] Ishan Chatterjee, Robert Xiao, and Chris Harrison. Gaze+gesture: Expressive, precise and targeted free-space interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ICMI '15, pages 131–138, New York, NY, USA, 2015. ACM.

[14] CIRS. Multi-purpose multi-tissue ultrasound phantom. `http://www.cirsinc.com/products/modality/67/multi-purpose-multi-tissue-ultrasound-phantom/`. [Online; accessed Oct 12, 2017].

[15] Yngve Dahl, Ole A Alsos, and Dag Svanæs. Fidelity considerations for simulation-based usability assessments of mobile ict for hospitals. *Intl. Journal of Human–Computer Interaction*, 26(5):445–476, 2010.

[16] Heiko Drewes and Albrecht Schmidt. Interacting with the computer using gaze gestures. *Human-Computer Interaction–INTERACT 2007*, pages 475–488, 2007.

[17] David Fono and Roel Vertegaal. Eyewindows: evaluation of eye-controlled zooming windows for focus selection. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 151–160. ACM, 2005.

[18] Classic Gaming. Space invaders online resources. `http://www.classicgaming.cc/classics/space-invaders/icons-and-fonts`. [Online; accessed Mar 1, 2017].

[19] Gazepoint. Gazepoint open gaze api by gazepoint. `http://www.gazept.com/dl/Gazepoint_API_v2.0.pdf`. [Online; accessed Oct 12, 2017].

[20] Gazepoint. Gp3 eye tracker. `https://www.gazept.com/product/gazepoint-gp3-eye-tracker/`. [Online; accessed Oct 12, 2017].

[21] Github. Eye movement data analysis toolkit (emdat). `https://github.com/ATUAV/EMDAT`. [Online; accessed May 29, 2017].

[22] Hartmut Glücker, Felix Raab, Florian Echtler, and Christian Wolff. Eyede: gaze-enhanced software development environments. In *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems*, pages 1555–1560. ACM, 2014.

[23] Joseph H Goldberg and Jonathan I Helfman. Comparing information graphics: a critical look at eye tracking. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond time and errors: novel evaLuation methods for Information Visualization*, pages 71–78. ACM, 2010.

[24] S.L. Hagen-Ansert. *Textbook of Diagnostic Sonography*. Number v. 1 in Textbook of Diagnostic Sonography. Elsevier/Mosby, 2012.

[25] Yasmin Halwani, Septimiu E. Salcudean, Victoria A. Lessoway, and Sidney S. Fels. Enhancing zoom and pan in ultrasound machines with a multimodal gaze-based interface. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 1648–1654, New York, NY, USA, 2017. ACM.

[26] Yasmin Halwani, Tim SE Salcudean, and Sidney S Fels. Multimodal interface design for ultrasound machines. In *Qatar Foundation Annual Research Conference Proceedings*, volume 2016, page ICTSP2476. HBKU Press Qatar, 2016.

[27] Dan Witzner Hansen, Henrik HT Skovsgaard, John Paulin Hansen, and Emilie Møllenbach. Noise tolerant selection by gaze-controlled pan and zoom in 3d. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, pages 205–212. ACM, 2008.

[28] Sandra G Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage Publications Sage CA: Los Angeles, CA, 2006.

[29] GE Healthcare. Logiq e9: Scan assistant. http://www3.gehealthcare.com/en/products/categories/ultrasound/logiq/logiq_e9/video_scan_assistant. [Online; accessed Oct 12, 2017].

[30] J Hempel, A Brychtova, Ioannis Giannopoulos, Sophie Stellmach, Raimund Dachselt, et al. Gaze and feet as additional input modalities for interacting with geospatial interfaces. 2016.

[31] Craig Hennessey. *Eye-gaze tracking with free head motion*. PhD thesis, 2005.

[32] Robert JK Jacob. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 11–18. ACM, 1990.

[33] Robert JK Jacob. Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in human-computer interaction*, 4:151–190, 1993.

[34] Robert JK Jacob. Eye tracking in advanced interface design. *Virtual environments and advanced interface design*, pages 258–288, 1995.

[35] Paul Kabbash, I Scott MacKenzie, and William Buxton. Human performance using computer input devices in the preferred and non-preferred hands. In *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*, pages 474–481. ACM, 1993.

[36] Yvonne Kammerer, Katharina Scheiter, and Wolfgang Beinhauer. Looking my way through the menu: the impact of menu design and multimodal input on gaze-based menu selection. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, pages 213–220. ACM, 2008.

[37] S Kardan, N FitzGerald, and C Conati. Eye movement data analysis toolkit (emdat) user manual. *University of British Columbia*, 2012.

[38] Dagmar Kern, Paul Marshall, and Albrecht Schmidt. Gazemarks: gaze-based visual placeholders to ease attention switching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2093–2102. ACM, 2010.

[39] Konstantin Klamka, Andreas Siegel, Stefan Vogt, Fabian Göbel, Sophie Stellmach, and Raimund Dachselt. Look + pedal: Hands-free navigation in zoomable information spaces through gaze-supported foot input.

In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ICMI '15, pages 123–130, New York, NY, USA, 2015. ACM.

[40] Nobuyuki Kobayashi, Eiji Tokunaga, Hiroaki Kimura, Yasufumi Hirakawa, Masaaki Ayabe, and Tatsuo Nakajima. An input widget framework for multi-modal and multi-device environments. In *Software Technologies for Future Embedded and Ubiquitous Systems, 2005. SEUS 2005. Third IEEE Workshop on*, pages 63–70. IEEE, 2005.

[41] Manu Kumar. Gaze-enhanced user interface design a dissertation submitted to the department of computer science and the committee on graduate studies of stanford university in partial fulfillment of the requirements for the degree of doctor of philosophy. 2007.

[42] Manu Kumar, Andreas Paepcke, and Terry Winograd. Eyepoint: practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 421–430. ACM, 2007.

[43] I Scott MacKenzie, Abigail Sellen, and William AS Buxton. A comparison of input devices in element pointing and dragging tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 161–166. ACM, 1991.

[44] Tobii Eye Tracker User Manual. Clearview analysis software copyright© tobii technology ab. 2006.

[45] Jennifer L Martin, Beverley J Norris, Elizabeth Murphy, and John A Crowe. Medical device development: The challenge for ergonomics. *Applied ergonomics*, 39(3):271–283, 2008.

[46] Emilie Mollenbach, Thorarinn Stefansson, and John Paulin Hansen. All eyes on the monitor: Gaze based interaction in zoomable, multi-scaled information-spaces. In *Proceedings of the 13th International Conference on Intelligent User Interfaces*, IUI '08, pages 373–376, New York, NY, USA, 2008. ACM.

[47] Catherine Plaisant, David Carr, and Ben Shneiderman. Image-browser taxonomy and guidelines for designers. *Ieee Software*, 12(2):21–32, 1995.

[48] Rahul Rajan, Ted Selker, and Ian Lane. Task load estimation and mediation using psycho-physiological measures. In *Proceedings of the*

21st International Conference on Intelligent User Interfaces, pages 48–59. ACM, 2016.

[49] Keith Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3):372, 1998.

[50] Andre Russo, Carmel Murphy, Victoria Lessoway, and Jonathan Berkowitz. The prevalence of musculoskeletal symptoms among british columbia sonographers. *Applied Ergonomics*, 33(5):385 – 393, 2002.

[51] Anthony Santella, Maneesh Agrawala, Doug DeCarlo, David Salesin, and Michael Cohen. Gaze-based interaction for semi-automatic photo cropping. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 771–780, New York, NY, USA, 2006. ACM.

[52] Anthony Santella, Maneesh Agrawala, Doug DeCarlo, David Salesin, and Michael Cohen. Gaze-based interaction for semi-automatic photo cropping. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 771–780. ACM, 2006.

[53] Selina Sharmin, Oleg Špakov, and Kari-Jouko Räihä. Reading on-screen text with gaze-based auto-scrolling. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*, pages 24–31. ACM, 2013.

[54] Sophie Stellmach and Raimund Dachselt. Investigating gaze-supported multimodal pan and zoom. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 357–360, New York, NY, USA, 2012. ACM.

[55] Sophie Stellmach and Raimund Dachselt. Look & touch: gaze-supported target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2981–2990. ACM, 2012.

[56] Sophie Stellmach, Sebastian Stober, Andreas Nürnberger, and Raimund Dachselt. Designing gaze-supported multimodal interactions for the exploration of large image collections. In *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*, NGCA '11, pages 1:1–1:8, New York, NY, USA, 2011. ACM.

[57] Veronica Sundstedt. Gazing at games: using eye tracking to control virtual characters. In *ACM SIGGRAPH 2010 Courses*, page 5. ACM, 2010.

[58] Yan Tan, Geoffrey Tien, Arthur E Kirkpatrick, Bruce B Forster, and M Stella Atkins. Evaluating eyegaze targeting to improve mouse pointing for radiology tasks. *Journal of digital imaging*, 24(1):96–106, 2011.

[59] Samuel Radiant Tatasurya. Multimodal graphical user interface for ultrasound machine control via da vinci surgeon console: design, development, and initial evaluation, Aug 2015.

[60] Dereck Toker, Cristina Conati, Ben Steichen, and Giuseppe Carenini. Individual user characteristics and information visualization: connecting the dots through eye tracking. In *proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 295–304. ACM, 2013.

[61] Irene Tong, Omid Mohareri, Samuel Tatasurya, Craig Hennessey, and Septimiu Salcudean. A retrofit eye gaze tracker for the da vinci and its integration in task execution using the da vinci research kit. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 2043–2050. IEEE, 2015.

[62] Ultrasonix. Touch screen ultrasound system. `http://www.ultrasonix.com/`. [Online; accessed Oct 12, 2017].

[63] Tamara van Gog, Liesbeth Kester, Fleurie Nievelstein, Bas Giesbers, and Fred Paas. Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior*, 25(2):325 – 331, 2009. Including the Special Issue: State of the Art Research into Cognitive Load Theory.

[64] F Vannetti, T Atzori, G Pasquini, L Forzoni, L Modi, and R Molino-Lova. Superficial electromyography and motion analysis technologies applied to ultrasound system user interface and probe ergonomics evaluation. *Advances in Human Aspects of Healthcare*, 3:227, 2014.

[65] Oleg Špakov. Comparison of eye movement filters used in hci. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 281–284, New York, NY, USA, 2012. ACM.

[66] Kimberly P Wynd, Hugh M Smith, Adam K Jacob, Laurence C Torsher, Sandra L Kopp, and James R Hebl. Ultrasound machine comparison: an evaluation of ergonomic design, data management, ease of use, and image quality. *Regional anesthesia and pain medicine*, 34(4):349–356, 2009.

[67] Shumin Zhai. What's in the eyes for attentive input. *Communications of the ACM*, 46(3):34–39, 2003.

[68] Shumin Zhai, Carlos Morimoto, and Steven Ihde. Manual and gaze input cascaded (magic) pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 246–253, New York, NY, USA, 1999. ACM.

[69] Dingyun Zhu, Tom Gedeon, and Ken Taylor. moving to the centre: A gaze-driven remote camera control for teleoperation. *Interacting with Computers*, 23(1):85–95, 2010.

Appendix A

Pixel-angle Accuracy Conversion for Eye Gaze Tracking Applications

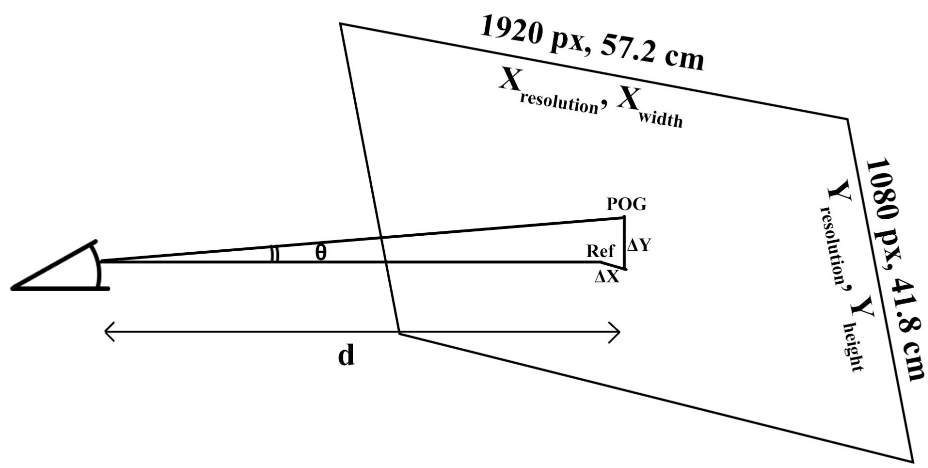


Figure A.1: Pixel-angle Conversion Parameters

In our user studies, we use Gazepoint eye gaze tracker, which has an angle error of 0.5 to 1 degrees. In this appendix, we calculate the error in pixels based on the apparatus used and explained in Chapters 5 and 6. We use the equations presented by Hennessy in [31]. An illustration of the parameters used for the calculations is shown in Figure A.1. A description of those parameters is the following:

- POG: Point of Gaze
- Ref: Reference Point
- ΔX: X Error
- ΔY: Y Error
- θ: Visual angle error = 1 degrees
- d: Distance To Screen = 60 cm

$$\Delta X_{cm} = \frac{57.2cm}{1920px} \Delta X_{px} = 0.0298 \Delta X_{px}$$

$$\Delta Y_{cm} = \frac{41.8cm}{1080px} \Delta Y_{px} = 0.0387 \Delta Y_{px}$$

$$\theta = 2 * \arctan \frac{\frac{\sqrt{X_{cm}^2 + Y_{cm}^2}}{2}}{d}$$

Assuming $\Delta X_{cm} \approx \Delta Y_{cm}$ and is represented by ΔE_{cm}

$$\theta = 1 = 2 * \arctan \frac{\frac{\sqrt{((\Delta E_{px})^2 (0.0298)^2 + (0.0387)^2)}}{2}}{60}$$

$$\tan \frac{1}{2} * 60 * 2 = \sqrt{((\Delta E_{px})^2 (0.0298)^2 + (0.0387)^2)}$$

$$65.5563 = \Delta E_{px} * \sqrt{(0.0298)^2 + (0.0387)^2)}$$

$$\Delta E_{px} = \frac{65.5563}{\sqrt{(0.0298)^2 + (0.0387)^2)}}$$

$$\Delta E_{px} = 21.4402px$$

Therefore, the error in pixels associated with the apparatus used in this thesis is a radius of 21.44 pixels.

Appendix B

Sonographers-Radiologists Survey

About the Survey

The aim of this survey is to understand ultrasound machine interface design from a user's perspective. We are currently working on studying how eye gaze trackers could improve interaction with ultrasound machines from a performance and ergonomic perspective. The questions relate to your daily use of ultrasound machines, musculoskeletal disorders due to work injuries, and current technologies used to improve the use of ultrasound machines. Your contribution and feedback are highly valued!

Section 1: General Information

1. Please select your current occupation
   - Sonographer
   - Radiologist
   - Cardiologist
   - Maternal Fetal Medicine Specialist
   - Student sonographer
   - Instructor
   - Other. Please specify: ......
2. In what age group are you?
   - 20 - 29
   - 30 - 39
   - 40 - 49
   - 50 - 59
   - 60 +
3. Gender:
   - Female
   - Male
4. Please specify your years of experience as a radiologist.
   - Not applicable
   - Less than 2
   - 2 - 5
   - 6 - 10
   - 11 - 20
   - 21 - 30
   - > 30
5. Sonographers: what types of ultrasound scans do you typically perform? (please select all that apply)

- General
- Cardiac
- Obstetric/Gynaecologic
- Vascular
- MSK

6. Radiologists: what types of interventional procedures do you typically perform? ......

7. Sonographers: on average, how long is the typical scanning time for an ultrasound scan?

   - $<$ 10 minutes
   - 10 - 20 minutes
   - 20 - 40 minutes
   - $>$ 40

8. Radiologists: on average, how long is the interventional procedure?

   - $<$ 10 minutes
   - 10 - 20 minutes
   - 20 - 40 minutes
   - $>$ 40

Section 2: Ergonomics and Work-Related Musculoskeletal Disorders (WRMSDs)

1. What is your typical work schedule?
    - Number of days per week: ......
    - Number of hours per day: ......
2. How often do you experience stress injuries or WRMSDs you believe is due to your career?
    - I have never experienced any WRMSDs or stress injuries due to my career.
    - Once or twice throughout my whole career.
    - Once every few years.
    - Once or twice per year.
    - Multiple times a year.
    - Continuously for ...... years.
3. How severe would you classify your WRMSDs?
    - Very painful and sometimes I have to take leaves from work.
    - Painful, but I can still manage the day.
    - Only a little distracting.
    - Intermittent
    - Not applicable.
4. What caused most of these injuries/disorders? Check all that applies.
    - Repetitive movement due to repetitive menu selection and button interactions.
    - Poor equipment design, such as height and location of the monitor, poor transducer grip design, chair height, patient bed location, etc.
    - Infrequent work breaks or short recovery time between scans.
    - Poor or awkward posture due to the type of scans performed.
    - Sustained force and pressure.
    - Other. Please specify: .....

Section 3: Efficient and Improved Ultrasound Interfaces

1. General Ultrasound Machine Usage and Familiarity

   (a) As an estimate, what is the percentage of settings that you frequently use in an ultrasound machine out of all the settings you are familiar with?
   - 10% - 30%
   - 30% - 60%
   - 60% +

   (b) Which buttons, functions or features do you use most frequently? (i.e. at least once per scan in > 90% of all scans)
   ......

   (c) Which buttons, functions or features do you use occasionally? (i.e. at least once per scan in 40% - 90% of all scans)
   ......

   (d) Which buttons, functions or features do you use rarely? (i.e. at least once per scan in < 40% of all scans)
   ......

   (e) Based on your own experience, if you were provided with a new ultrasound machine with a slightly different interface (in terms of the layout of the buttons and/or the software interface) than the one you are used to and have been using for most of your work, how long do you think it would take you to find your way around the different settings you often use in your scans/procedures?
   i. Less than one working day (I can find my way around a new ultrasound machine very easily)
   ii. About a few days to a week (I can find my way around a new ultrasound machine easily, but I face some struggles sometimes)
   iii. More than a week (I find it really hard to adjust to a new ultrasound machine interface)

   (f) Please provide your rating for the following, where 1 = Highly Disagree, 7 = Highly Agree.
   i. Scanning anatomical structures that are in constant motion can be a time-sensitive task that requires an efficient and responsive user interface.
   1 - 2 - 3 - 4 - 5 - 6 - 7

ii. Ultrasound foot-switches can be helpful in repetitive tasks (such as freeze or print)
1 - 2 - 3 - 4 - 5 - 6 - 7

iii. Sometimes I have to go through a lot of steps (through interface menus) to select a particular setting.
1 - 2 - 3 - 4 - 5 - 6 - 7
If you selected 5 or higher, please provide examples: ......

iv. I switch my attention between the monitor and the ultrasound interface buttons very often and it gets distracting sometimes
1 - 2 - 3 - 4 - 5 - 6 - 7

v. I switch my attention between the monitor and the ultrasound interface buttons very often and it makes me lose focus of important image details sometimes
1 - 2 - 3 - 4 - 5 - 6 - 7

(g) Please describe the hardware or software interface features in some ultrasound machines you worked with which you think are very efficient and ergonomically convenient in comparison to other ultrasound machines. (E.g. the keyboard is located on the same panel as the rest of the buttons, the touch screen provides quick suggestions for annotations, patient data is loaded automatically, etc.) ......

(h) Please describe the hardware or software interface features in some ultrasound machines you worked with which you think are NOT efficient and not ergonomic in comparison to other ultrasound machines. (E.g. keys are distributed in a non-intuitive way, using the touch screen frequently is distracting, etc.) ......

2. Evaluation of Existing Improvements to Ultrasound Machine Interfaces

(a) Have you used automated ultrasound exam software such as the "Scan Assistant" that is available with some GE machines?
   - Yes
   - No (skip to question (e))

(b) What are the types of scans that you typically use this type of software for? ......

(c) From your experience, what is the added benefit from using this software? (Check all that apply)

- Significantly shortens the duration of an ultrasound exam.
- Contributes to lowering the risk of WRMSDs as it lowers the amount of repetitive movements with physical ultrasound machine interface buttons.
- Helps me not to forget any steps for a particular ultrasound exam and organizes my work flow
- Helps me focus more on the patient
- Other: ......

(d) Can you think of any drawbacks of automating the ultrasound exam steps using software like Scan Assistant? .....

(e) Have you used any voice-enabled ultrasound machines? For example, the speech recognition feature in Philips iU22.

- Yes. Could you please describe your experience and why (or why not) you would prefer to use this feature during your scans or procedures: ......
- No.

3. Hands-free Interaction with Ultrasound Machines (Radiologists)

(a) Do any of your scans or procedures require a sterile environment or sterile equipment?

- Yes. Examples? ......
- No

(b) Is there a need for adjusting ultrasound image settings and parameters during an interventional procedure? If yes, what are the typical settings changed?

- Yes. All types of procedures frequently require it.
- Yes, but the frequency of this need changes based on the type of interventional procedure being performed
- No, most of the procedures have image settings pre-set. There might be exceptions though.
- No, not at all.
- Other. Specify: ......

(c) Are there any cases where an assistant (e.g. another sonographer) is required for your scans or procedures? (If yes, continue. If no, skip the rest of this section)

- Yes

- No (skip to the next section)

(d) Are there any cases when it is difficult to communicate intent to an assistant?

- No, instructions are very straightforward.
- It depends on the experience and background of the assistant.
- Yes, but it's tolerable and does not affect the flow of the procedure.
- Could you please provide examples? ......
- Yes, it would be much easier if I could change the parameters directly.
- Could you please provide examples? ......

(e) Would having some hands-free control method to the ultrasound machine reduce your need for an assistant?

- Yes, significantly.
- Yes, but the assistant might still know better in terms of ultrasound machine settings control.
- No, I do not prefer to interact with the ultrasound machine at all.

Section 4: Efficient and Improved Ultrasound Interfaces

Do you have any other observations, comments. thoughts, suggestions, or ideas you'd like to offer on these topics that might be of use in our research? ......

Appendix C

General Survey Feedback from Sonographers

This appendix lists down all the answers received to the last question in our sonographers survey, as discussed in Chapter 3: "*Do you have any other observations, comments. thoughts, suggestions, or ideas you'd like to offer on these topics that might be of use in our research?*".

- "*It's definitely time for a major change! I think voice recognition command with a headset would be very beneficial, in addition to reinventing the foot pedal for freeze and cine and finding ways to compress data and simply record exams rather than freezing and printing images for normal exams.*"
- "*Gel warmer on right side.*"
- "*Toggles are easier to work with than knobs.*"
- "*Probe design that is lightweight, wireless if possible and fits into hands well.*"
- "*Does the monitor have to be attached to the machine? For some exams, a ceiling suspended monitor may be more convenient.*"
- "*Ultrasound exams by their nature do not follow the same order. Multiple methods need to be looked at in order to address the wide spectrum of exams. Scan assist, foot switches, and voice command have all been used. Each of these technologies have uses and drawbacks. Simplified keyboards with multi-use keys have been used on several systems. This method has not been very effective as most technologist prefer the ability to see every available parameter. Single optimization buttons have helped but cannot completely replace the technologist's ability to change parameters to suit every patient body type. Effective layout of a keyboard with the ability to adjust based on body type of the technologist is the best option. Areas currently needing consideration are VR goggles,*

cordless transducers and robotic exoskeletons. These technologies are currently being developed or researched. In the development of these new technologies, engineers need to consider the interaction between technologist and the patient. Important information is always discovered during the casual conversations during the exam. Care must be taken in order to keep barriers to a minimum. Adoption of new methods has always been problematic. Technologists as a rule develop their own personal routines for exams. These routines ensure that nothing is missed during the exam. New technologies should be developed to enhance their ability while allowing technologists to maintain their current work flow. This could improve the adoption of new technologies."

- "*It will be great with hands free typing program like voice recognition.*"
- "*The buttons shouldn't be hard to press. The panel of the machine should be able to adjust with ease. The wheels on the machine should be smooth and easy to steer. The screen can be tilted to any angle.*"
- "*Voice recognition would be useful for interventional work. the ability of the system to know when you would want to make an adjustment such as depth of field, focus or overall gain. The image optimization button does this to somewhat of extent using voice for printing instead of a button. The image should not be blurry as it is sometimes. Storage of video clips via voice (especially fetal echoes).*"
- "*Voice recognition might be nice, but only if you can get a patient to be quiet. We do children and infants, so not sure how that would work.*"
- "*It certainly sounds very interesting. Again, I would like to see a feature of the machine that either measure circumference either by touch screen, or auto measure (like NT measurement package on the GE feature).*"
- "*The machines need to cater to taller sonographers. It would be nice to get the machine closer to the bed. Also more efficient design of keys on the console would be helpful.*"
- "*It is difficult when a machine is not conducive to being moved to either side of a bed. Doppler of the arm is an example when this is needed, or ICU portables. Machines need to have more similar interfaces so that one can easily move from one type of machine to another.*"

- "*Reduce the size of the machine. Make sure all user parts are wireless (transducer, monitor to cart, ...etc.) Ask the manufacturers to really leverage computer and mobile phone technologies in designing ultrasound machines. There is no need for 100kg machines, or transducers that need a "fishing pole" to hold the cable.*"

- "*A machine that has a pull out/retractable area for your feet to rest on. Even better ability to scan large patients without using all of our own arm strength. I have already had to have RTC surgery due to my work.*"

- "*The machines need to be smaller (or maybe they are, but my hospital doesn't have the money to get them) and lighter. It would be nice to squeeze a machine into an ICU room between the 2 IV poles, the vent, and all the other equipment that is on the side of the bed that I want my machine on. A lighter machine that can be moved with one hand would help with every DVT study where you start at the groin, but need to move all the way to the calf. Good luck with your research, I hope we (sonographers) see the benefits in the years to come.*"

- "*Probe cords are heavy: a cordless probe would be nice but only if the probe itself was as light and small as it is right now.*"

- "*I am 6'2 The machines tend not to go high enough for me to stand at a comfortably level.*"

- "*I have had trouble with Lt elbow pain. I relate to extending my arm for annotations, Rt arm/neck pain from pressing with probe. If the machine can penetrate with less force on obese patients it would help. Ergonomically, the ports need to be easier to access and more of them so you do not need to bend and change as often. Lighter body for moving and locking mechanisms that are sturdy. And maybe machine/bed combo that has a cut out bed, in order to be closer to patients and have the keyboard so you are looking ahead instead of arms spread and body turned. Good luck with that!*"

Appendix D

First-iteration Clinical User Study

In this appendix, we present our user study with sonographers during the first iteration of our user-centred design.
To evaluate the designed systems, we designed a controlled user study to compare both of the design alternatives. The aim of the study is to observe the time on task and interaction repetitiveness between the base system (the manual-based system) and the gaze-supported system and to receive qualitative feedback from sonographers on the introduced interaction that will guide us for the second design iteration.

D.1 Gaze-supported Interface Design

The interface design is the same as detailed in System Design (Chapter 4) as *alternative 1*.
In this alternative, it is important to note that the user is not receiving any visual feedback of where their current point of gaze is within the image for the One-step Zoom alternative. The only feedback that the user is receiving regarding their eye gaze is whether or not the eye tracker has lost the user's eyes through two red circles shown at the bottom of the interface, as shown in Figure D.1 (a). This feedback is also implemented with the second alternative.

D.2 Apparatus

As for the hardware, we used the Gazepoint GP3 eye gaze tracker [20] with the accompanying Open Gaze API. The system was implemented and displayed on the screen of the Ultrasonix touch [62] machine. The ultrasound image was streamed through Ulterius API. The software was written in Python and the interface implemented with PyQt.

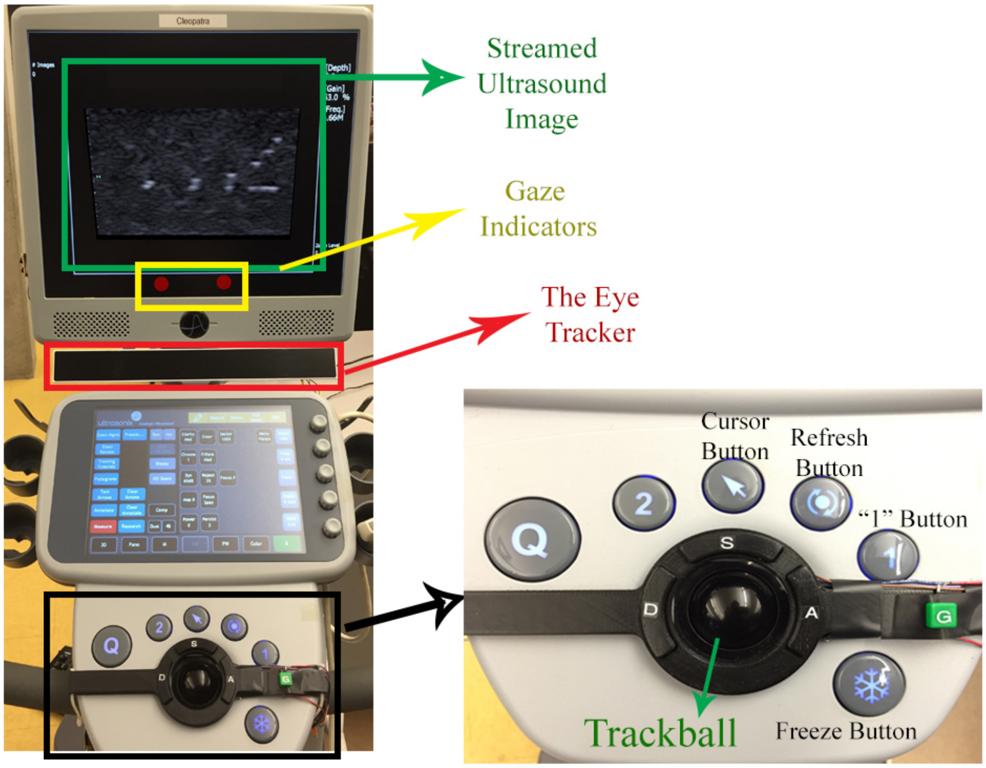


Figure D.1: The Ultrasound Machine and System Components Used for the Implementation of Our Systems and User Study (Left). The Control Keys Panel of the Ultrasound Machine Used in Our Users Study (Right).

It is important to note that this ultrasound machine is not the typical type of machine used in diagnostic sonography, but is often used in point of care and by non-sonographers due to its simple touch-based interface with a few physical buttons and limited capabilities. To resemble the multipurpose buttons interface of the ultrasound machines typically used in diagnostic sonography, we prototyped and 3D-printed a set of buttons around the trackball of the Ultrasonix Touch machine that captured these characteristics, as shown in Figure D.1. Figure D.1 also shows the complete setup and the position of the eye tracker relative to the ultrasound machine monitor.

The functionalities of the buttons are slightly different between OZ and MZ. Table D.1 shows the mapped functions to each button on the ultrasound controls panel.

Table D.1: The Dedicated Buttons' Functionalities Based on the System Used

System	**Button**	**Functionality**
One-step Zoom (OZ)	A	Zoom in
	D	Zoom out
Multi-step Zoom (MZ)	A	Zoom in / out
	D	Hold
	Trackball	Resize
OZ and MZ	S	Scroll (Pan)
Ultrasonix System	G	Capture Image
	A, D, S, G	Inactive
	Trackball	Move / resize
	Cursor button	Toggle between Move /resize
	Refresh button	Zoom in / zoom out
	1	Capture Image
All Systems	Freeze button	Freeze image

D.3 Procedure

A discussion with each participant followed the experiment to receive their general impression of the new eye tracking-based system and identify its strengths and drawbacks given their experience and background in sonography.

A total of 6 participants, 1 male and 5 females, took part in the user study. Four of the participants completed the user study wearing glasses and one completed the user study wearing contact lenses. The participants were all, except one student sonographer, expert sonographers who perform either obstetric exams, general exams or both.

For simplicity, and due to the availability and the evidence from previous literature [66] on the use of phantoms for ultrasound machine interface usability evaluation, the tasks were performed with a CIRS [14] quality assurance phantom. A phantom is a specially-designed object used in medical training in place of a patient made of material that mimics real tissues. It contains several targets that can be acquired with US imaging. For each different system tested, the participant sonographer was first given the time they needed to get familiar with the interface by freely

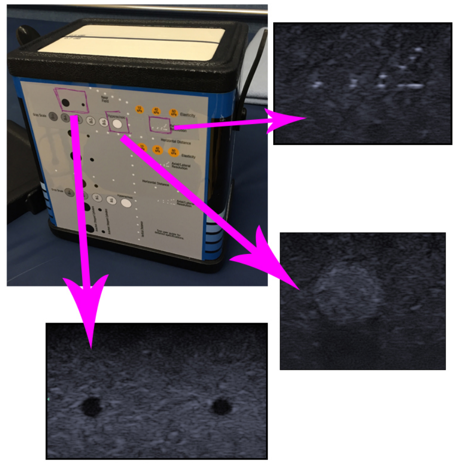


Figure D.2: The Quality Assurance Phantom Used in the User Study and the Corresponding Ultrasound Images of the First, Second and Third Targets.

exploring the phantom and acquiring and zooming into targets. Afterwards, the participant was asked to zoom into and capture three predefined targets, as shown in Figure D.2.

The participant could change the image parameters related to TGC and gain, but only before starting the zoom acquisition tasks, so that the zoom interaction could be isolated from other ultrasound machine-related interactions for later analysis.

Each sonographer carried out all three tasks on all three systems. The first system the participants used was the conventional ultrasound machine interface, which we will refer to as the "base system". This acts as a baseline for comparisons with subsequent interactions with OZ and MZ.

The second system alternated between OZ and MZ: half of the participants tested OZ as the second system and MZ as the third and the other half vice versa.

D.4 Results

For each system, the average completion time and button hits, including the use of the trackball, of the three tasks was observed.

D.4.1 Time on Task

We observed that the average task completion time using the base system is the shortest ($M = 14.45$, $SD = 5.72$) followed by OZ ($M = 21.15$, $SD = 12.1$) and the longest is MZ ($M = 40.65$, $SD = 18.36$). Higher time periods associated with the gaze-supported alternatives could be due to the participants' unfamiliarity with the system when using the multi-modal gaze-based interface.

D.4.2 Button Hit Rate

Provided the nature of the eye gaze-supported design, we observed that the rate of button presses using the base system is the highest ($M = 0.51$, $SD = 0.12$) followed by OZ ($M = 0.3$, $SD = 0.07$) and the lowest is MZ ($M = 0.25$, $SD = 0.05$). To provide a deeper understanding of the input rate, table D.2 shows the average of the all the tasks for the number of buttons hit, completion time and input rate for each of the systems tested.

Table D.2: The averages of participants' tasks results for the number of buttons hit, completion time and input rate for each of the systems tested.

Technique	**Button hits**	**Time (s)**	**Input rate**
Base	7.83	15.45	0.51
OZ	5.72	21.15	0.27
MZ	8.78	40.65	0.23

D.4.3 Qualitative Feedback and Discussions

In terms of qualitative feedback received from the participants, they have shown their preference to OZ over MZ. P3 stated:

"I liked OZ the most of out the three. I found it to be the least visually-distracting compared to MZ. When we are scanning, we are doing a lot of visual assessment of the tissue itself. Extra overlays that take us away from seeing tissue pathology might be a negative distraction."

Latching the zoom box movement to the eye movement, even with filtering and workarounds to reduce the distraction factor as explained in the design section, was still perceived as highly distracting by all participants. Another phenomenon we observed is the participants' struggle to perfectly place the window around the target before confirming the zoom action, despite the fact that they can later refine the location through the panning feature. Although the zoom box is present in the base system, it moves only in direct response to trackball movement, which is intentionally initiated by the user; thus, no participants reported distraction with the base system.

In terms of panning, there has been a variety of feedback regarding the usefulness of the feature. The majority of the sonographers found it very helpful as it reduces the need to adjust the probe to move the image once the image is zoomed in. On the other hand, other sonographers found it a little unnecessary since they prefer using the probe to move the image around.

As for the time measurements, we notice very high standard deviation figures for all interfaces, which suggests that there might be other factors influencing the time taken to complete a task that we did not take into account, such as the user's adaptability or tiredness.

One of the interestingly observed behaviours of the participants during the user study is their change in posture when using the gaze-based systems in comparison to the base system. Participants seemed more aware of their posture to keep their head within the field of view of the eye tracker. Some participants did not glance at the keys at all, as they did not want the gaze tracker to lose their gaze while glancing elsewhere. This is an example of an unnatural behaviour resulting from using eye tracking systems, which requires gaze trackers with higher field of view or interface enhancement that is less sensitive to movements. Nevertheless, P4 found that as a positive result of using gaze-based systems, which can implicitly alert the sonographer to always stay in an upright posture to avoid occupational back injuries.

As for qualitative feedback, participants preferred simpler gaze-based systems that do not have any visual feedback of where they are looking. In other words, they prefer to "trust the eye tracker" to determine where they are looking and not provide distracting feedback surrounding areas of interest.

D.5 Improvements for the Second Iteration

We observed several issues with the design, user study structure and evaluation from our first iteration. In our second design and evaluation iteration, we target to improve the following challenges we faced:

Apparatus Providing the participants with an ultrasound machine interface with full capabilities has its advantages and drawbacks. The advantage is designing a more realistic user study environment and getting more realistic results of the user interaction. However, a drawback, which outweighed the advantage at this stage of research, is the participants' distraction from the intended gaze-supported interaction we are testing: the participant sonographers spent some of the user study and task time interacting with other ultrasound features to adjust the image parameters, which took away from their focus on the zoom feature being evaluated. One way to overcome this challenge is limiting the physical interface to a number of ultrasound functions we are interested in testing, which we design in our second design iteration.
Additionally, the physical interface layout of the base interface is different from the physical interface layout of the gaze-supported systems: we added extra buttons that are not normally part of the ultrasound machine interface and that added some confusion to the participants.
We also decide to remove in our second iteration the gaze indicators, as we found that participants rarely used them and mainly relied on instructions from the researcher regarding their optimal posture for eye gaze detection, given the little amount of time they are allowed to interact with and their unfamiliarity with the system.

User Study Structure Involving the end users in the design and evaluation iterations is essential in the user-centred design we adopt. However, we realize that we might be able to better identify and isolate the gaze-supported interaction effects by running evaluations separate from the context of sonography. This is due to the discussed cognitive load factors in sonography including concurrent image analysis and patient interaction, which could mask the real effects of the newly-introduced interaction design. In our second iteration, we design a user study that targets lay users and abstracts the interaction away from sonography. The premise is that if the abstract gaze-supported interaction yields positive results with lay users, then running a user study with sonographers has potential in an improved interaction. Otherwise, if the gaze-supported interaction in a simple abstract setting is not successful, then it will not have great potential with the added sonography-related cognitive load. Another issue with the user study structure is the sample size and the number of targets acquired per participant. We acknowledge that the analyzed data set is too small and that our statistical analysis is at a very preliminary stage that does not allow for reliable statistical conclusions yet.

User Study Tasks One important issue with our user study design is the lack for defined and uniform training tasks across all participants. Allowing the participants to interact with the system "as long as they need to learn it" is not a useful instruction for users of a system for the first time. We observed that participants committed multiple interaction-related errors while performing the tasks. These errors could be significantly reduced if we define training tasks sufficient for the participants to follow and learn the system interaction.

Evaluation Although we measure the button hit rate, we do not take into account the amount of time a button is held down, which is another form of physical interaction. Additionally, repetitive interaction is not a suitable measure, as other issues rise with low repetitiveness, such as unnatural postures and extended focus on the gaze interaction, we suspect another type of mental cognitive load rising as the physical repetitiveness decreases.

Appendix E

Game User Study Script

E.1 Participant Recruitment Email

Hello awesome participant!
Thank you for your interest in participating in the user study!
Before I sign you up, I have to first make sure you are eligible to participate. Please reply back to this email with answers to the following questions:

1. Do you wear glasses?
2. Do you wear bifocal/gradual glasses (the ones for both far-sighted and short-sighted vision correction)
3. Do you have any abnormal (whether diagnosed or undiagnosed) eye condition? (e.g. lazy eye after fatigue)
4. Do you have any left arm/hand/fingers injury? Do you have any pain associated with the movement of your left arm/hand/fingers for any reason?
5. Do you have any previous experience with operating ultrasound machines? (operating the machine itself, not being the patient)

Please select a time from the available times below. It is recommended that you choose a time within the period of your best performance during the day, if possible, as you will be actively learning and applying new simple tasks (no previous knowledge required).
Time Schedule: https://doodle.com/poll/gh7eke9hgexk232k
The consent form describes the experiment procedure of the general larger project, which is aimed at enhancing the interaction of clinicians with ultrasound machines. The experiment procedure for this user study is essentially the same, but designed for non-clinicians with no background of operating ultrasound machines (therefore it is an interactive game instead). Attached is your copy of the consent form for your reference.

Thank you again for dedicating the time to play video games for science and research!

E.2 OZ Preparation Settings

- Disable the zoom knob press + gaze button press
- Python constants: GRAPHICAL_EXPERIMENT = 1
- Run g_nob_fastZoom.py

E.3 MZ Preparation Settings

- Disable the zoom knob rotations
- Python constants: GRAPHICAL_EXPERIMENT = 2
- Run g_nob_detailedZoom.py

E.4 Before the Participant's Arrival

1. Raise the participant's monitor to a viewable level.
2. Switch the researcher's monitor away (to not distract the participant while using the system).
3. Place the controls box in front of the participant + make sure it's stable.
4. Hardware: connect the eye tracker and the controls box.
5. Prepare the consent form, evaluation form, reward form, researcher's form and participant reward.
6. Block any direct light source to the eye gaze tracker.
7. Run the following software:
   (a) Gazepoint Analysis
   (b) Gazepoint Control
   (c) Windows Media Player (for the game sound tracks)

   (d) Pycharm: and run the python program of the particular experiment (Make sure LOG_RESULTS_FLAG is set to True)
   (e) Delete all previous temporary data from the previous participant.

E.5 After the Participant's Arrival

1. Provide consent form
2. Provide participation reward
3. Provide demographics survey
4. Check if eye tracker can see eyes and can calibrate!
5. Start timer
6. Introduce user study "you are here to play space invaders!" + introduce structure (2 sessions: each has 2 games. After each session you will fill out the following form)
7. Present evaluation form + give the participant a minute to read the TLX reference + mention that the participant is not being personally evaluated and all data will be kept confidential
8. Adjust seating in front of screen:
   (a) Left hand can reach controls
   (b) Eye tracker can see eyes
   (c) Ask the participant to clean their glasses if wearing ones
   (d) Seat height is comfortable
9. Test calibration
10. Test volume
11. Start screen capture
12. Turn off lights above experiment area + close curtains
13. Prepare laptop with word document for discussion + any other notes during the sessions

E.6 Manual-based Interaction Session

1. Run demo
    (a) Software:
        i. Explain the game
        ii. Show targets
        iii. Show timer
        iv. Show level
        v. Show level progress
        vi. Timeout results in: losing a life + repeating the whole level
        vii. Shooting and scoring results in: level progress bar going up
    (b) Hardware:
        i. Zoom in and out
        ii. Trackball to pan
        iii. Red button to shoot
            A. "You can shoot monster only when it turns purple"
            B. "It turns purple only when "
                - "The full alien is within the view"
                - "The alien is at a particular zoom level"
    (c) Demo 3 aliens
    (d) Other instructions:
        i. "You are only allowed to use your left hand"
        ii. "I will not respond to any questions or comments during your gaming sessions"
        iii. "If you have any questions about the game design, you can ask after the experiment is done"
2. Re-run program for the training session and the recorded session
    (a) Run soundtrack
    (b) Calibrate
    (c) Run training
    (d) When done, pause soundtrack + go to next soundtrack
    (e) Run soundtrack
    (f) Calibrate

   (g) Run recorded
   (h) When done, pause soundtrack + copy log + go to next soundtrack

3. Provide participant with evaluation form and mention: “only consider the last game you played for these questions. E.g. the temporal demand of the last game”.

E.7 Break Session

Offer the participant some chocolate and ask the participant to relax for a few minutes and let the researcher know when they are ready for the next session.

E.8 Gaze-supported Interaction Session

1. Run program for gaze demo
   (a) Calibrate
   (b) Software: Exactly same as the manual-based interaction session
   (c) Hardware
      i. Turn on gaze mode
      ii. “this is using the position of your eye gaze to specify where to zoom”
      iii. “turn the knob and see what happens”
      iv. “you might run into situations where eye gaze is inconvenient or not fast enough to select the target, you can still use the trackball”
      v. If you move your head, it won’t help, although it might feel natural to do so
   (d) Demo 3 aliens
2. Re-run program for training and recorded
   (a) Calibrate
   (b) Run soundtrack
   (c) Calibrate
   (d) Run training
   (e) When done, pause soundtrack + go to next soundtrack

   (f) Run soundtrack
   (g) Calibrate
   (h) Run recorded
   (i) When done, pause soundtrack + copy log

3. Provide participant with evaluation form and mention: "only consider the last game you played for these questions. E.g. the temporal demand of the last game".

E.9 Discussion Session

1. Double check forms after participant fills them out (look for any missing data)

2. "Do you have any general questions / comments?"

3. "For each system, mention 3 advantages and 3 disadvantages (record voice + type)"

4. "Please elaborate on your sources of frustration (from the evaluation form)"

5. "Please keep the contents of the user study confidential until the end of the data collection phase."

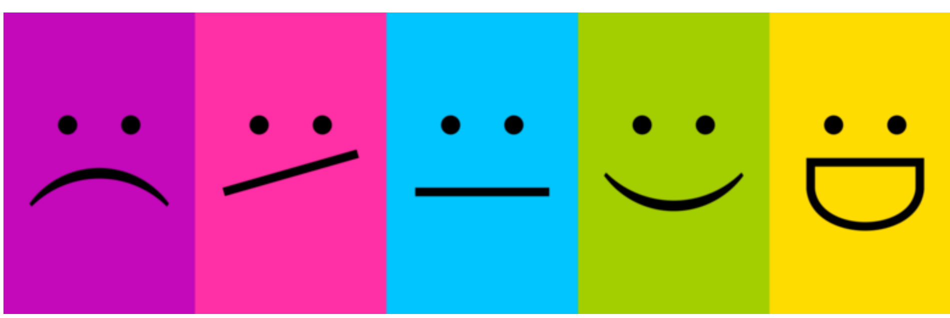
Appendix F

Game Participants' Demographics Form

1. Please specify your age range.
   - 18 - 25
   - 26 - 35
   - 36 - 45
   - 46 - 55
   - 56 - 65
   - Above 65
2. You classify yourself as:
   - Right-handed
   - Left-handed
   - Ambidextrous
3. What is your laptop's default touch-pad scroll direction settings? (please double check with the researcher if you are unsure how to respond to this question)
   - Natural scroll (like Mac computers)
   - Reverse scroll (like most Windows-based computers)
4. Do you have any current or past conditions with your vision, including past surgeries? (please double check with the researcher if you are unsure how to respond to this question)
   - Yes
     - Shortsightedness. Level: ......
     - Astigmatism

  - Vision correction surgery. Type: ......
  - Other: ......
- No

5. Have you used an eye tracker before?
   - Yes. Please provide details, including your frequency of use: ......
   - No
6. Are you a frequent user of 2D/3D image editing and design software that offer panning and zooming images? (e.g. Photoshop, SolidWorks, ... etc.)
   - Yes
     - Please mention the name(s) of the software: ......
     - What is your frequency of usage
       * Daily
       * Weekly
       * Monthly
       * Occasionally
   - No
7. Which of the following best describes your mental tiredness level right now?



Appendix G

Game Qualitative Evaluation Form

The following table is provided to the participant as a reference.

[**Mental Demand**] Did the game require a lot of thinking?
[**Physical Demand**] Was the game physically easy or demanding, physically slack or strenuous?
[**Temporal Demand**] How much time pressure did you feel due to the pace at which the game occurred? Was the pace slow or rapid?
[**Overall Performance**] How successful were you at winning the game? How satisfied were you with your performance? NOTE: this is your OWN performance, NOT the system's performance*
[**Frustration Level**] How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the game?
[**Effort**] How hard did you have to work to win the game?

Note Two copies of this form are provided per participant to evaluate each session independently (manual-based interaction and gaze-supported interaction)

---

1. Sources of Load (Weights)
   Circle the member of each pair that contributed more to the workload of the task. (E.g. were you more stressed by X or Y?)

Effort Performance	Temporal Demand Frustration
Temporal Demand Effort	Physical Demand Frustration
Performance Frustration	Physical Demand Temporal Demand
Physical Demand Performance	Temporal Demand Mental Demand
Frustration Effort	Performance Mental Demand
Performance Temporal Demand	Mental Demand Effort
Mental Demand Physical Demand	Effort Physical Demand
Frustration Mental Demand	

2. Magnitude of Load (Ratings)

   Rate each of the following based on the task performed

   - Mental Demand
     Low _ High
   - Physical Demand
     Low _ High
   - Temporal Demand
     Low _ High
   - Performance
     Low _ High

- Effort
  Low _ High
- Frustration
  Low _ High

3. What were your main sources of frustration? (select all that applies)

   - Game difficulty due to time limit
   - Game difficulty due to limited lives
   - Aliens don't turn purple easily
   - Moving the trackball
   - Physical interface layout (the control box)
   - Eye gaze tracking accuracy
   - Other: ....

4. Rate the following statements

   (a) I thought the system was easy to use
   Strongly disagree _ _ _ _ _ Strongly agree

   (b) I think that I would need the support of a technical person to be able to use this system (if you were to learn this system on your own)
   Strongly disagree _ _ _ _ _ Strongly agree

   (c) I found the various functions in this system were well integrated
   Strongly disagree _ _ _ _ _ Strongly agree

   (d) I thought there was too much inconsistency in this system (E.g. things do not perform as you expect sometimes)
   Strongly disagree _ _ _ _ _ Strongly agree

   (e) I would imagine that most people would learn to use this system very quickly
   Strongly disagree _ _ _ _ _ Strongly agree

(f) I found the system very cumbersome (difficult) to use
Strongly disagree _ _ _ _ _ Strongly agree

(g) I felt very confident using the system
Strongly disagree _ _ _ _ _ Strongly agree

(h) I needed to learn a lot of things before I could get going with this system
Strongly disagree _ _ _ _ _ Strongly agree

Appendix H

CBD-CHD Ultrasound Scan Steps

Below are the typical steps followed by sonographers to perform general Common Bile Duct (CBD) and Common Hepatic Duct (CHD) exams, based on an observation and discussions with two sonographers.

1. Ask the patient to lie in supine position.
2. Choose the abdominal pre-set on the ultrasound machine, which loads the default settings and parameters for abdominal scans (such as dynamic range, frame rate, frequency, gain, etc.).
3. Pick up the curvilinear probe and apply warm gel on it before scanning the patient.
4. Locate the CBD within the abdominal area.
5. Since the depth is set to 13 cm by default (on the particular machine that was being observed), decrease the depth (to about 9 cm) to capture the area closer to the probe in higher definition.
6. Interchangeably, change the gain to obtain a brighter image.
7. Once located, decrease the depth further and zoom into the area including the common hepatic and bile ducts.
8. Move the zoom window until the target is located at the centre of the image.
9. Place the focal zone at the level of the CBD and CHD.
10. Freeze the image.
11. Select the measurement calipers.

12. Place the first caliper at the first edge of the structure to be measured, and the second caliper at the second edge.
13. Measure the CHD, then repeat steps 11 and 12 to measure the CBD (or vice versa).
14. To get better images and measurements, sometimes the patient is asked to change their position to lie on their left side.
15. The same steps are repeated from 4 to 13 and new images and measurements are obtained of the same structures.

Appendix I

Clinical User Study Script

I.1 Participant Recruitment Email

To include in the email:

1. Location of the user study (attach a map)
2. A contact phone number (the lead researcher's)
3. Parking information and instructions
4. Instructions to reach the lab
5. "Please plan to arrive about 15 minutes before your scheduled time just in case of any difficulties finding the location"
6. "If you wear prescription glasses and own contact lenses, it is recommended to wear contact lenses for this experiment as it does not take as long to calibrate with the eye gaze tracker"
7. "A reward of a $10 Starbucks card is offered for participation"

I.2 Before the Participant's Arrival

To bring to the experiment area:

1. Two monitors
2. Ultrasound machine
3. Interface controls box
4. Screw driver (for the controls box, in case its height needs to be adjusted)
5. Printouts of participant forms

6. Printout of user study structure
7. Printout of phantom targets
8. An iPad with target ultrasound images

To place on the table:

1. User study structure
2. Gift card
3. Gift card sheet
4. Consent form
5. Demographics form
6. Researcher's form

Technical setup:

1. Turn on ultrasound machine and Ultrasonix software at least 30 minutes before user study session to allow it to load
2. Connect ultrasound machine with an Ethernet cable to the computer running the client software
3. On ultrasound machine, run Pycharm and SERVER_main.py
4. Organize the patient room
5. Prepare towels
6. Put away old used towels
7. Turn off lights above experiment area (add a sticky note to both sides of the lab that a user study is currently running)
8. Add a sign at the patient room that a user study is in progress, please do not disturb
9. computer (client) setup:
    (a) Setup monitor extended display setting

   (b) Run us_machine_client_interface.py (every time this script stops, the server script has to be restarted for this client script to run again)

10. Place the printout targets over the ultrasound machine monitor

11. Place the iPad of the target ultrasound image in front of the participant

12. Set the ultrasound machine parameters:

    (a) Attach the correct probe (Curvilinear C5-2)
    (b) Preset: abdomen - general
    (c) Zoom: 100%
    (d) Freq: 2.5M
    (e) Focus (at the end of the image)
    (f) Depth: 14.0cm
    (g) Gain: 74%

13. Delete (or move) all previous data in the temporary results folder

I.3 Introduction Session

1. The researcher introduces herself

2. Provide consent form

3. Provide participation reward

4. Provide demographics survey

5. Start screen capture recording

6. Start timer

7. Introduce the project

    (a) "My project is the integration of eye gaze tracking with the interaction with ultrasound machines."
    (b) "I only worked on the zoom functions: Low-resolution zoom + hi-resolution zoom"

8. "Since we are focusing on the interaction, the zoom features are all digital zoom-based functions, so the produced image is never as high quality as the acquired images with the typical high-res zoom function in ultrasound machines."
9. Adjust seating
10. Test calibration
11. Start audio recording (if needed)
12. Demo the software
    - (a) The ultrasound image
    - (b) The context view
    - (c) The image capture area
    - (d) The available tools
    - (e) The "recording" indicator
13. Demo the hardware
    - (a) "This is designed to match the layout as closely as possible to typical ultrasound machines used in sonography"
    - (b) "Try to use zoom on your own using both zoom techniques ... "
    - (c) The rest of the knobs work as well, as you are familiar with them, the implementation is not comprehensive as I'm mainly focusing on the interaction with the zoom function.
14. Demo the gaze-supported features
15. Mention the following gaze-supported interaction instructions
    - "For the Multi-step Zoom technique, it is best to use the eye gaze feature only when moving the zoom box for long distances across the image".
    - "Eye gaze is very jittery in small areas, therefore, it is best to use the trackball for fine motions of the zoom box around the target".
16. List further instructions...

- "If you need to change any settings of the image (other than the ones available in front of you), you can do that only during the demo session and before starting the training sessions. Please do not reach to change any image settings during the phantom or patient tasks if not absolutely necessary.".
- "Keep in mind that this user study takes an hour to complete, so if you have any particularly long feedback during the experiment, please leave it to the discussion session so we can finish all tasks on time".

17. Phantom exploration:
    (a) "For the phantom, I would like you to scan only the edge area of the 0.5 dB part of the phantom".
    (b) Show phantom targets
    (c) Place each group of targets one by one on the ultrasound machine's screen.

I.4 Phantom Session

The same instructions in the introduction follow in the phantom session, only with a different set of targets. The phantom session is further divided into the gaze-supported interaction sub-session and the manual-based interaction sub-session.

I.5 Patient Session

The lead researcher volunteers to be scanned by the sonographer. Instructions given to the sonographer include:

1. "Perform the CBD scan with the default zoom setting and level you normally use in your scans".
2. "For the first 5 image captures, use the gaze-supported interaction alternatives of your choice. For the second 5 image captures, use the manual-based interaction alternatives of your choice".
3. "Between each image capture, please lift your right hand (holding the probe) and place it again to re-locate the CBD.".

I.6 Discussion Session

The following are the list questions discussed with each participant sonographer:

1. What type of ultrasound machine are you familiar with? Did this hardware/software layout and functions closely resemble the ultrasound machines you typically use in your ultrasound scans?
2. What type of zoom do you typically use in your scans? And when? And how frequently?
3. What kind of shapes do you normally need to zoom into?
4. Are they regular shapes with defined centres?
5. How would you describe your own perceived eye gaze behaviour when zooming into targets in an ultrasound image? Do you mainly focus on the target or do you keep peripherally scanning the rest of the image?
6. What advantages and disadvantages did you find using the proposed eye gaze-supported zoom system?

Appendix J

Sonographers' Demographics Form

1. Please specify your age range.
   - 18 - 25
   - 26 - 35
   - 36 - 45
   - 46 - 55
   - 56 - 65
   - Above 65
2. Please specify your years of experience as a sonographer.
   - Not applicable
   - Less than 2
   - 2 - 5
   - 6 - 10
   - 11 - 20
   - Above 30
3. You classify yourself as:
   - Right-handed
   - Left-handed
   - Ambidextrous
4. Do you have any current or past conditions with your vision, including past surgeries? (please double check with the researcher if you are unsure how to respond to this question)

- Yes
  - Please provide details: .....
- No

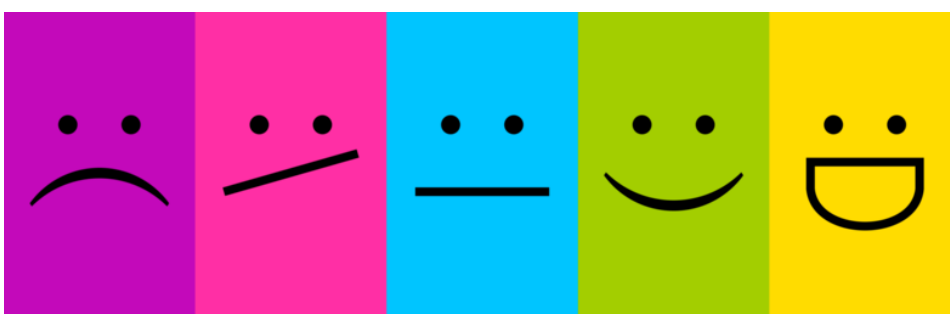
5. Have you used an eye tracker before?

   - Yes. Please provide details, including your frequency of use: ......
   - No

6. What types of ultrasound scans do you typically perform? (Please select all that applies)

   - General
   - Cardiac
   - Obstetric/Gynaecologic
   - Vascular
   - MSK
   - Other. Please specify: ......

7. Which of the following best describes your mental tiredness level right now?



Appendix K

Post-processing of the Collected Data from the Clinical User Study

Table K.1: Summary on Clinical User Study Data Post-Processing

Family	**Measure**	**Data Transformation Applied**	**#Trimmed Outliers**	**Violated Assumptions After Post-processing**
Sum	Time	Log10	4	None
	Number of Fixations	Log10	5	None
	Absolute Path Angles	Log10	5	None
	Fixation Duration	Log10	6	None
	Path Distance	Log10	9	Input method (Manual): Kurtosis = 1.11, K-S Normality = 0
				Technique (MZ): Skewness = -1.11, Kurtosis = 3.02
	Relative Path Angles	Log10	5	None
Rate	Absolute Path Angles	None	3	None

Family	Measure	Data Transformation Applied	#Trimmed Outliers	Violated Assumptions After Post-processing
Rate	Eye Movement Velocity	None	3	Technique (Mixed): Kurtosis = -1.82
	Fixations	None	5	Technique (Mixed): Kurtosis = -2.31
	Relative Path Angles	None	2	Technique (Mixed): Kurtosis = -1.78
Mean	Absolute Path Angles	None	7	Technique (MZ): Kurtosis = 1.66
	Fixation Duration	Log10	6	None
	Path Distance	None	3	None
	Relative Path Angles	None	5	None
Standard Deviation	Absolute Path Angles	None	3	Technique (Mixed): Kurtosis = -1.55
	Fixation Duration	Log10	5	Technique (Mixed): Kurtosis = -2.316
	Path Distance	None	3	Technique (MZ): Kurtosis = 2.0
	Relative Path Angles	None	5	Technique (Mixed): Kurtosis = -1.5