Enhancing Zoom and Pan in Ultrasound Machines with a Multimodal Gaze-based Interface

Yasmin Halwani
University of British Columbia
Vancouver, BC, Canada
yhalwani@ece.ubc.ca

Victoria A. Lessoway
British Columbia Women's
Hospital and Health Centre
Vancouver, BC, Canada
vickie@lessoway.ca

Septimiu E. Salcudean
University of British Columbia
Vancouver, BC, Canada
tims@ece.ubc.ca

Sidney S. Fels
University of British Columbia
Vancouver, BC, Canada
ssfels@ece.ubc.ca

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053174

Abstract
We present the first iteration of a user-centred design for integrating an eye gaze tracker with ultrasound (US) machine interfaces used in routine diagnostic sonography. The aim is to speed up the interaction of the sonographer with the machine, offer a more ergonomic solution while minimizing the cognitive load and maintaining the produced image quality. In this iteration, we target pan and zoom functions. Field studies and observations have led to two design alternatives. A feasibility study of two design approaches determined from field studies is done using a functional prototype. Results from six sonographers provided evidence for the potential of a multimodal gaze-based interface for US machines. Results from the feasibility study have also led to a second design iteration that combines the advantages of both gaze-based designs.

Author Keywords
multimodal interaction; gaze input; eye tracking; gaze-based interaction; zoom; pan; sonography; ultrasound machines

ACM Classification Keywords
H.5.2 [User Interfaces]: Input devices and strategies

Introduction
Sonographers spend hours of daily work acquiring and modifying parameters of images. Those images are sent

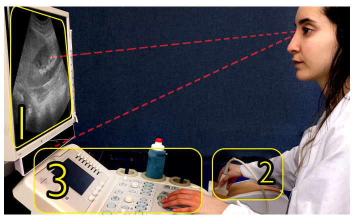

Figure 1: A diagnostic sonographer's contexts of attention: the US image, the manual keys, and the patient

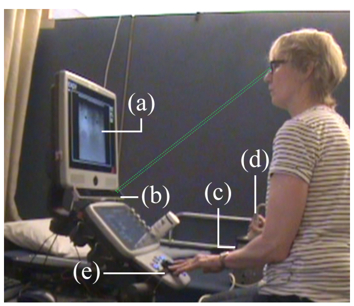

Figure 2: The user study setup. (a) the ultrasound image (b) the gaze tracker (c) the phantom (d) the ultrasound probe (e) the manual inputs

later to physicians for further review and diagnosis. US machine interfaces are designed with efficiency in mind to make the access to US functions as fast as possible; as every minute counts toward the throughput of US exams per day and the perceived healthcare quality by the patient.

A study on the prevalence of musculoskeletal disorders among British Columbia sonographers [8] 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 that was distributed to a local ultrasonographers' society revealed that nearly half of the respondents (N = 48) reported repetitive movements due to menu selection and physical keys interaction as a major cause of their experienced occupational musculoskeletal injuries.

As shown in Figure 1, US machines require substantial bimanual interaction; as the sonographer works with the US probe in their right hand, while manipulating the console with their left (normally, even left-hand dominant sonographers are trained for this configuration). Due to concurrent bimanual interaction, patient communication and image analysis; such an interaction requires a higher cognitive load than typical desktop environments. Introducing modifications to the US machine interface should not increase the cognitive load a sonographer has to already deal with. Instead, modifications are expected to simplify the interaction and therefore reduce some of this load.

We adopt existing interaction techniques of gaze-supported zoom and pan systems. Our aim is to evaluate those techniques within the context of sonography. We start from basic techniques found in the literature, as will be explained in the next section, and build gradually to more complex ones.

Related Work

Recent advances in gaze-supported applications research show that multimodal gaze-based interaction has potential in improving speed and user satisfaction. As an example, Mollenbach et al. [7] experimentally showed an improvement of task performance for pan and zoom over traditional input methods.

Other studies showed that multimodal gaze-based pan and zoom integrated with mouse input followed closely to the pure mouse interaction in the aspect of speed, spatial awareness and accuracy, even with the lack of prolonged user exposure to gaze-supported interfaces. As an example, the study presented by Stellmach et al. [9] is an application for Google maps. Similarly, the work presented in [1] investigates the same application with different interactions. In our work, we adopt their Dual-to-Zoom approach, which combines gaze with manual input buttons.

In addition to efficiency, Zhai et al. [11] argues that enhancing interfaces with gaze trackers has the potential to also reduce repetitive stress injuries for computer users. Recent research [4] also leveraged gaze tracking to reduce repetitive stress injuries in different application areas.

Field Study

US machines, functions, and exam durations differ based on the specific anatomy being imaged during an US exam. Given such diversity, we need to get a practical view of these factors and study the feasibility of integrating a gaze tracking system with US machines. Seven different types of US exams were observed at two different hospitals. Surveys to a local sonographers' society and structured interviews and informal discussions with two sonographers were continuously conducted to bring in professionals' perspectives. The results dictated the next steps and helped us

Zooming and panning require the sonographer to perform a number of steps before the image is zoomed. Once the probe is positioned over the required area, the following steps are performed by the sonographer:

1. Enable the “zoom mode”,
2. Using the trackball, move the zoom box to the location of the ROI 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.

select one function to start from that will benefit the most from a multimodal interface.

Some sonographers reported that the large amount of options sometimes cause unwanted distraction, which draws attention away from the US image. A common scenario in obstetrics, for example, 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.

We observed that the zoom function causes repetitive physical interaction with the machine interface when a sonographer repeatedly and manually alternates between selecting and resizing particular areas in the image to be magnified. On the other hand, the literature shows that zoom interfaces has the potential to benefit from integrating a gaze tracker and thus we focused our attention to this particular US function.

Sonographers assess the acquired image from a holistic perspective and, in most cases, only one object of interest is present at a time to zoom into, such as a gallbladder surrounded by other organs, a tumour surrounded by tissue, or a fetal heart surrounded by the fetal organs. The purpose of performing pan and zoom is to obtain a higher resolution image of the area of interest to perform further accurate functions, such as taking measurements.

We have identified the risks of deploying a gaze-based interaction technique based on observed user behaviour, US machine capabilities and the clinical environment setting. Although the lighting conditions in an US exam room is optimal for gaze tracking, calibration is always required for high accuracy gaze tracking. Given the typical length of a general US exam of at least 20 minutes, the routine changes in sonographer positions and the frequent context switches, there is a risk of gaze tracker calibration deterioration.

System Design

We test two design alternatives to zoom into and acquire US images: Simple Zoom (SimZ) and Resizable Zoom (ResZ). The first one implements the basic case of zooming into a fixed-size region of interest (ROI) surrounding the point of gaze (POG). The second is a more consistent version with the design of the high-resolution zoom feature in US machines: it provides the user with a higher level of control over the dimensions of the ROI.

For both designs, pan regions are defined at the visible boundaries of a zoomed image. If the POG falls within one of these regions while simultaneously holding the pan manual trigger, as shown in figure 4, the image scrolls in that direction. Additionally, whenever the POG is invalid (e.g. when the user is outside the tracker’s field of view), the user receives visual feedback from the system.

Design Alternative 1: Simple Zoom (SimZ)

As illustrated in Figure 5, this alternative uses one manual input to zoom into a ROI, and a separate manual input to zoom out. The first “zoom-in” action (120%) is performed based on the location of the POG and all consecutive “zoom-in” actions center the ROI further by a factor of 30%. “Zoom-out” backtracks until the original image is restored.

The dimensions of the ROI are constant and are proportional to the dimensions of the original image. In this design, the user is not provided with a visual feedback of the location of the POG prior to performing a zoom action.

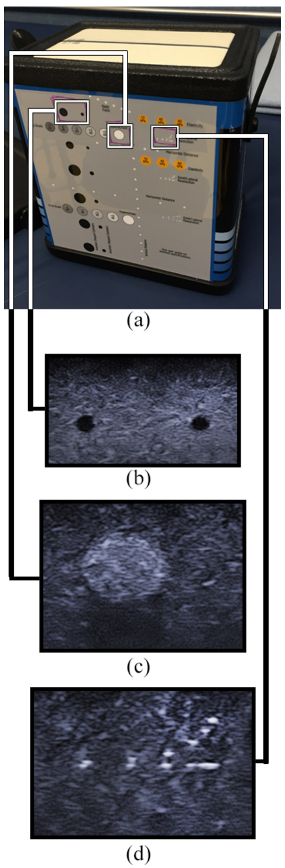

Figure 3: (a) The quality assurance phantom used in the user study. The side of the box shows illustrations of targets that are visible when imaged with ultrasound. (b), (c) and (d) are the corresponding ultrasound images of the first, second and third targets.

Design Alternative 2: Resizable Zoom (ResZ)
Contrary to the first alternative, this design provides the user with a visual feedback of the user's POG. The ROI is represented as a resizable "zoom box", as shown in Figure 6. The zoom box is latched to the POG and is resized through the US machine's trackball.

Once the dimensions of the box are set and the location of the area of interest is locked in position, the user initiates a manual trigger to confirm the zoom action. The same manual trigger is used later to reset the image to its original state.

A simple averaging filter is implemented 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. Once the user starts resizing the box, the box stops following the eye movement to allow for precision. However, one unavoidable risk that the interaction could run into is the deterioration of calibration due to a resulting positive feedback loop, as explained in classic literature on gaze tracking interfaces [5].

Evaluation of Prototype

The aim of this initial evaluation of the system's feasibility is to provide greater insight into the technical advantages and drawbacks of the system within the context of sonography, which guides the next iteration of the interface design.

Apparatus
For both design alternatives, we use the same apparatus and tools. As for the hardware, we use the Gazepoint GP3 gaze tracker [3] with the accompanying Open Gaze API. The system was implemented on the Ultrasonix touch [10] machine. The US image was streamed through Ulterius API. The software was written in Python and the interface implemented with PyQt. The US machine used in this study is the Ultrasonix Touch. Figure 2 shows the complete setup and the position of the tracker relative to the US machine monitor.

We use the default 5-points calibration that comes with the tracker. The amount of time each user spent using the tracker, and therefore the frequency of head movements, did not call for re-calibrations.

For simplicity, the tasks are performed on a CIRS [2] quality assurance phantom instead of a patient. 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, as shown in Figure 3.

Procedure
A total of five professional practising sonographers, who perform either obstetric exams, general exams or both, and one student sonographer performed a set of US imaging tasks that comprised of zooming into and capturing targets using both design alternatives. Four of the participants completed the user study wearing glasses and one completed the user study wearing contact lenses. Four of the participants never interacted with a gaze tracker before, while the rest had a previous interaction with gaze trackers. Calibration was performed prior to using each of the design alternatives using the default 5-point calibration of the Gazepoint GP3 system.

We focused on qualitative feedback given that our targeted users are a selected group of professionals, which we do not have access to in large numbers at this stage of research. In addition, with such a small number, it will be hard to interpret quantitative results to represent the whole user group. A discussion with each participant followed the

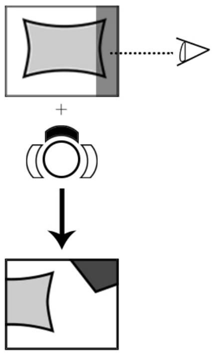

Figure 4: Panning in both gaze-based alternative designs

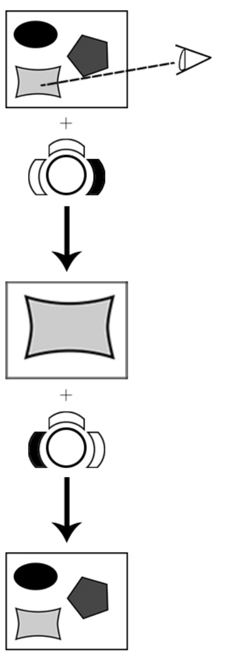

Figure 5: Simple Zoom (SimZ)

experiment to receive their general impression of the new gaze-based system, and to identify its strengths and weaknesses, given their experience and background in sonography.

For each different system tested, the participant sonographer was first given a few minutes to get familiar with the interface by freely exploring the phantom and zooming into targets. Afterwards, the participant was asked to zoom into and capture three predefined targets.

Each sonographer carried out all three tasks on all three systems. The first system the participants used was the conventional US machine interface, which we will refer to as the "base system". This acts as a baseline for comparisons with subsequent interactions with SimZ and ResZ. The second and third systems were counterbalanced.

Results and Discussion

Time and Ergonomics

We observed that SimZ requires the lowest amount of interaction with the US machine's manual inputs among the other two systems, and consequently less time to finish the required tasks. Thus, we see a potential in using a multimodal gaze-based interaction with US machines to reduce occupational stress-injuries due to repetitive physical interaction with the machine's controls, only if a suitable interaction design approach is followed.

One of the interesting 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 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 gaze 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.

Cognitive Load

Through discussions, participants have shown their preference to SimZ over ResZ. P3 stated:

"I liked SimZ the most out of the three. I found it to be the least visually-distracting compared to ResZ. 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 box movement to the POG, even with filtering and workarounds to reduce the distraction factor, 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, which could add unwanted cognitive load to the task. Participants preferred simple gaze-based systems that do not present visual feedback of their POG. In other words, they prefer to "trust the gaze tracker" to determine the position without distracting visuals surrounding the ROI.

In terms of panning, there has been a variety of feedback regarding the usefulness of the feature. Four sonographers found it very helpful as it reduces the need to adjust the probe to move the image. On the other hand, two sonographers preferred bimanually using the probe to move the image around while optimizing its parameters.

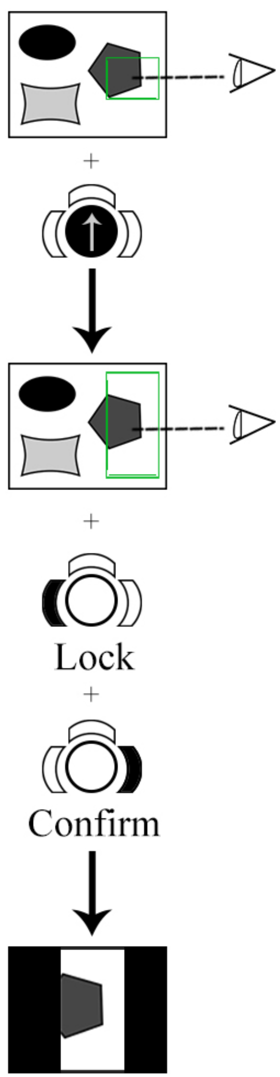

Figure 6: Re-sizable Zoom (ResZ): ROI always moves with the POG, unless the user locks it in place

Second Iteration Design

Testing our proposed systems with practicing sonographers brings insight into the advantages and potential drawbacks of applying gaze-supported interfaces in the context of sonography. The major advantage that the SimZ design brings is its manual input simplicity. Also, it does not impose any visual distractions on the US image, which does not contribute to higher cognitive load. On the other hand, the zoom box available in the ResZ design provides the user with the ability to set the dimensions of the acquired zoomed US image, which is important to isolate distractions in an US image and only keep the target anatomy in the image. One disadvantage of the SimZ design is the limited control available to the user as the zoom ratio and ROI dimensions are fixed. As for ResZ, the latched movement of the zoom box to the point of gaze adds extra cognitive overhead, and reduced learnability by the participants.

Inspired by the work presented in [6] and our preliminary results, we propose a new design (Combined Zoom: ComZ) that combines the best of both design alternatives for our next refinement. In this design, the user looks at the ROI, initiates a trigger by pressing and holding the zoom-in input, a zoom box appears for further dimensions modifications with the trackball, and finally releases the button to confirm the zoom action. This approach engages the sonographer's muscle groups to hold the mode temporarily which mitigates potential mode errors. Therefore, the visual overlay is available only upon request and does not cause any positive feedback loops and calibration degradation.

Conclusion and Future Work

Earlier work in the area of gaze-supported pan and zoom interfaces show a promising improvement over traditional interfaces for some applications. A sonographer's US interface requires substantial bimanual operation with visual cues of where the sonographer's attention is. The high cost of sonographer time, as well as the common hazardous strain injury, suggest this application may benefit substantially from an eye-gaze interface. To investigate, we completed the first iteration of our design-prototype-evaluate cycle. We assess the feasibility of two different gaze-based zooming and panning techniques (SimZ and ResZ) applied to US that exploits sonographers' task space that involves visual search targeting one object within a noisy scene.

From our initial feasibility study, we are optimistic that eye-gaze can be added effectively to US interfaces to help speed up the interaction and lower workplace injuries for sonographers, without adding unwanted cognitive load. By abstracting the interaction, our results do not solely apply to integrating gaze tracking with US machines, but to any other interface/machine in an environment with higher cognitive load introduced by concurrently performing other bimanual and analysis tasks. Our results in this short paper serve as a stepping-stone to pursuing further studies in this direction.

In future work, a more elaborate measure should be used to assess the ergonomic factors of the proposed new design. We are currently looking at such measures for our follow-up user study, which is to be designed to quantitatively analyze the effectiveness of the refined US design in terms of speeding up the image acquisition process and the interface ergonomics. Moreover, solutions on how to compensate for calibration deterioration due to the reasons observed in the field study should be further explored for more effective and deployable systems.

Acknowledgements
Many thanks to British Columbia Ultrasonographers' Society, BC Women's Hospital, Richmond Hospital, and Gazepoint for their continuous help and support.

References
[1] Nicholas Adams, Mark Witkowski, and Robert Spence. 2008. The Inspection of Very Large Images by Eye-gaze Control. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '08)*. ACM, New York, NY, USA, 111–118. DOI: http://dx.doi.org/10.1145/1385569.1385589

[2] CIRS. 2016. Multi-purpose multi-tissue ultrasound phantom. (2016). http://www.cirsinc.com/products/modality/67/multi-purpose-multi-tissue-ultrasound-phantom/.

[3] Gazepoint. 2016. GP3 Eye Tracker. (2016). https://www.gazept.com/.

[4] Hartmut Glücker, Felix Raab, Florian Echtler, and Christian Wolff. 2014. EyeDE: Gaze-enhanced Software Development Environments. In *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 1555–1560. DOI:http://dx.doi.org/10.1145/2559206.2581217

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

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

[7] Emilie Mollenbach, Thorarinn Stefansson, and John Paulin Hansen. 2008. 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)*. ACM, New York, NY, USA, 373–376. DOI: http://dx.doi.org/10.1145/1378773.1378833

[8] Andre Russo, Carmel Murphy, Victoria Lessoway, and Jonathan Berkowitz. 2002. The prevalence of musculoskeletal symptoms among British Columbia sonographers. *Applied Ergonomics* 33, 5 (2002), 385 – 393. DOI:http://dx.doi.org/10.1016/S0003-6870(02)00038-8

[9] Sophie Stellmach and Raimund Dachselt. 2012. Investigating Gaze-supported Multimodal Pan and Zoom. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 357–360. DOI:http://dx.doi.org/10.1145/2168556.2168636

[10] Ultrasonix. 2016. Touch screen ultrasound system. (2016). http://www.ultrasonix.com/node/73.

[11] Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 246–253. DOI: http://dx.doi.org/10.1145/302979.303053