Virtual Reality for User-Centered Design and Evaluation of Touch-free Interaction Techniques for Navigating Medical Images in the Operating Room

Anke Reinschluessel1,3
anke.reinschluessel@tzi.de

Joern Teuber2,3
jteuber@cs.uni-bremen.de

Marc Herrlich1,3,5
herrlich@eit.uni-kl.de

Jeffrey Bissel4
Melanie van Eikeren4
Johannes Ganser4
Felicia Koeller4
Fenja Kollasch4
Thomas Mildner4
Luca Raimondo4
Lars Reisig4
Marc Ruedel4
Danny Thieme4
Tobias Vahl4
ivor_pro@uni-bremen.de

Gabriel Zachmann2,3
zach@cs.uni-bremen.de

Rainer Malaka1,3
malaka@tzi.de

1Digital Media Lab
University of Bremen, Germany
2Computer Graphics Group
University of Bremen, Germany
3Creative Unit: Intra-operative information
University of Bremen
4University of Bremen, Germany
5University of Kaiserslautern
Kaiserslautern, Germany

Abstract
Computer-assisted surgery has pervaded the operating room (OR). While display and imaging technologies advance rapidly, keyboard and mouse are still the dominant input devices, even though they cause sterility problems. We present an interactive virtual operating room (IVOR), intended as a tool to develop and study interaction methods for the OR, and two novel touch-free interaction techniques using hand and foot gestures. All was developed and evaluated with 20 surgeons. The results show that our techniques can be used with minimal learning time and no significant differences regarding completion time and usability compared to the control condition relying on verbal instruction of an assistant. Furthermore, IVOR as a tool was well received by the surgeons, although they had no prior experience with virtual reality. This confirms IVOR is an effective tool for user-centered design and evaluation, providing a portable, yet realistic substitution for a real OR for early evaluations.

Author Keywords
touch-free interaction; medical imaging; operating room; virtual reality

ACM Classification Keywords
H.5.1 [Information interfaces and presentation (e.g., HCI)]: Multimedia Information Systems – Artificial, augmented,
and virtual realities; H.5.2 [Information interfaces and presentation (e.g., HCI)]; User Interfaces – Input devices and strategies (e.g., mouse, touchscreen)

Introduction
Medical imaging and computer-assisted surgery (CAS) have pervaded the modern operating room (OR). One aspect and the primary impulse behind their success has been providing novel solutions to difficult or unsolved medical problems, e.g., by visualizing structures that were not visible in the past. However, leveraging the full potential of such systems for medical procedures in general and especially for day to day procedures hinges on how well they can be integrated and utilized by physicians and other medical experts inside and outside the OR. Therefore, developing efficient and reliable interaction techniques and input devices for medical image data are key areas that need improvement.

This paper presents a novel approach to apply modern user-centered design and development methodologies (UCD) to develop and evaluate interaction techniques and user interfaces for browsing medical image data inside the OR. While we developed and evaluated our approach using a specific and relevant task (browsing medical images) as a test-bed, the approach is not limited to this task and can be applied to all kinds of human-computer interaction (HCI) problems inside the OR.

Our approach makes use of the recent advances in virtual reality (VR) technology to build a virtual test environment that integrates well with the traditional stages of UCD, i.e., idea generation, prototyping, and iterative development and user studies. Employing VR technology in addition to established methods greatly facilitates the iterative development process and the collection of empirical data with expert users, e.g., surgeons, in a situated context, without the limits, dangers, and logistic challenges of early evaluation inside the OR.

In this paper, we present the Interactive Virtual Operating Room (IVOR). Additionally, we provide two interaction methods, hand and foot gestures, for image browsing, which are integrated into IVOR for evaluation. We discuss design decisions, technical aspects, and the results of the user study with regard to the iterative and user-centered design process. Our findings show that the new techniques were successfully used by the surgeons even with limited training time. While others have suggested similar techniques, there is a lack of studies comparing the different techniques against each other in a situated setting.

Related Work
VR is an option to do (medical) research, training and procedures [25]. Seymour’s study demonstrated that skills that have been acquired in a VR training setting can be transferred to a clinical or animal laboratory setting [23]. Seymour et al. [24] also showed that training in VR transfers technical skills to the OR environment. Additionally, the use of VR surgical simulation significantly improved the OR performance during laparoscopic interventions. Grantcharov et al. [8] use a randomized double-blind trial to produce evidence that objectively assessed intra-operative errors of laparoscopic interventions can be significantly reduced by VR training. Gallagher et al. [6] showed that “VR is more likely to be successful if it is systematically integrated into a well-thought-out education and training program which objectively assesses technical skills improvement proximate to the learning experience”. VR elements were incorporated in medical training simulations by Billinghurst et al. [5] and by von Zadow et al. [28]. One build a simulator to train paranasal sinus surgery procedures [5] whereas
the other developed SimMed, a tool to train medical students on a simulated, life-sized patient displayed on a large touch-screen [28]. King et al. [12] used VR to view medical images and showed that there is no significant difference in accuracy.

User centered design [4] approaches are already been used in the medical environment [14]. Their research group developed a general work station in the 90ies using UCD with medical practitioners [20].

Figure 1: Schematic overview of the IVOR hardware and software setup. We use two infrared tracking cameras controlled by the tracking PC. The tracking data is sent via ethernet and the VRPN protocol to the VR PC, which runs Unreal with plugins, the IVOR and displays it on an HMD. A Leap Motion controller is connected to the VR PC and used via the corresponding Unreal plugin.

Browsing images is an important task and often used to test new OR systems [16, 26]. In order to provide surgeons with a more efficient, comfortable, precise, and sterile interaction technique, the feet and hands can be an effective means of accomplishing this goal in comparison to other modalities, such as voice or gaze interaction. Touch-less gesture interaction [9, 11, 13, 17, 19, 27, 29] is an option to interact with imaging systems, displays, and controllers without breaking the sterility barrier. 42 % of the surgeons intuitively use both hands for gestures capable to browse through images [1]. Graetzel et al. [9] were the first, to our knowledge, who implemented a gesture recognition system for the OR. They used the system as a mouse replacement to control a computer with simple point-and-touch gesture. Intentionally, the system should be used in minimal invasive procedures. Grange et al. [7] also used detection and tracking of the hand to substitute the mouse and interact with the interface with two “click” gestures. Others were using wristband inertial sensors and capacitive floor sensors to detect foot and hand movements [10]. The hand gestures were used for browsing, the foot gestures for enabling or disabling the interaction system. The tap gesture was more intuitive to the participants than the swipe gesture. The authors came to the conclusion, that their foot gesture system might be a valid alternative for other interaction techniques.

Interactive Virtual Operating Room

To meet the requirements of motion tracking and high-resolution virtual environment (VE) we used state-of-the-art hard- and software. The HTC Vive head-mounted display (HMD) was used in combination with the Leap Motion to provide room-scale VR with hand-tracking. Therefore we were able to show the hands inside the VR and to use them for gesture control. For the foot control, we used an optical tracking system and modified OR shoes with reflectors. The tracking system is set up on a separate PC which sends the tracking information via Ethernet using a VRPN server-client system (see Figure 1). On the software side, we used the Unreal Engine with plug-ins for the Leap Motion and VRPN to provide a realistic rendering of the VE and integrate the novel interaction methods. The VE was created by using an existing, realistic model of an OR and modifying it according to our needs.

This model of the OR was designed based on input from surgeons and surgical staff. During the first iterative design cycles the users noted that the room did not feel interactive enough and was too empty as ORs are usually quite crowded. Therefore we modeled the area around the operating table to be more like it is mid-intervention, with tables with surgical tools, an open patient with abdominal surgical site and covered in the usual green cloth. We placed several monitors in the virtual space which can display the slices of a set of CT images. These displays were augmented with control displays for the foot and hand interaction. The resulting model of the OR can be seen in Figure 2. The users described the virtual OR as realistic and overall agreed that it is a great tool to visualize new ideas and that it is possible to use it for user studies.
Touch-free Interaction Design

Visits to the OR and interviews with the surgeons revealed that interacting with the image data during the surgery is a crucial task and browsing medical images is of key interest to the medical staff. Given the hygiene constraints and the current WIMP (Window Icons Menu Pointer) interfaces the surgeons cannot directly interact with the images. The surgeons have to verbally instruct non-sterile OR staff to interact with the images on their behalf, e.g., by using commands like “up” or “down”. This “UpDown” procedure is currently state of the art but error-prone and high-demanding for staff members in terms of technical skills. Furthermore, the surgeon is not in direct control of the actions performed.

Our observations showed that during retrieving information from the images, the surgeons have at least one hand or the feet available as an input method if not already used for input in form of buttons on the floor [21]. Previous research on interaction methods in the OR confirms this observation (e.g., [18], [22], [10], [9], [29], [15]). As hand and feet are available for interacting with the system, we developed two approaches: a single hand-based interaction method and a foot-based interaction method. A detailed description of the gestures is given in the next paragraphs.

A bar on the display provides feedback for both input methods (see Figure 3 & 5). It shows whether the hand or feet are detected and which browsing speed is chosen. The image in Figure 5 shows that both feet are detected and the highlighted “pause” symbol indicates that no browsing action is active. The symbols for “pause”, “play”, “fast forward”, and “fast backward” are based on a media player metaphor. The interface provides additional feedback for the foot position on the floor. The hand interface is designed likewise. In Figure 3 the left hand was detected and is currently activating the fast forward paging down mode.

For hand-based interaction different gesture sets have been developed for intra-operative use. Whereas some use pointing gestures [15] [9], others use hand gestures [22] [10], which involve the different finger positions and yet others incorporate more space-consuming gestures that involve lower arm movement [29] [9]. They all support a variety of image manipulation functions and require a sophisticated hand recognition. Based on the results of our observations and interviews we focused on a solution with an easy and clear hand gesture. Following the approaches of Mewes et al. [18] and Wachs et al. [29] we are using a roll gesture for browsing through the images. Their results state that a “circling” gesture is “easy, precise, [and] robust” and the mapping to the slicing steps of the images is useful [18].

Feet are an integrated input method in the OR as the surgeons are using foot pedals. These fit the hygiene constraints in the OR but Allaf et al. [3] showed that finding the pedals is a source of distraction in the OR. The approaches of Jalaliniya et al. [10] and Alexander et al. [2] avoid pedals by using foot gestures. Based on their research, the interviews with the medical staff, and feedback from HCI experts we decided to develop a set of foot gestures to browse image data using discreet and continuous gestures involving both feet.

The implemented gestures work similar: one hand, held upright, is raised into the field of detection of the Leap Motion Controller. The feet are standing straight for the null position. Rotating the hand or lifting, moving and lowering the right foot clockwise is paging down. The mirrored gesture performs the paging up action. The feet have an additional “tap gesture” for single image browsing as in lifting and lowering either feet. The lifting is necessary as our observations and pre-tests showed that the properties of hospital floors prevent easy feet rotation with shoes. These gestures
supports the browsing behavior, we observed during our hospital visits. The surgeons tend to skim through the images and immediately stop when they found what they were looking for.

**User Study**

The user study was conducted in comparable conference rooms at different hospitals. The apparatus used in the setup is listed in the margin on the left. After collecting their informed consent, the participants were randomly assigned to one of the interaction methods to start with: hand, foot or UpDown (baseline condition). The instructor explained the interaction method to the participant. Following this, the participant entered IVOR and had some time to get comfortable in the VE. The participant browsed a training stack of images before the two test stacks started. The task was to find a specific artifact we included randomly in the stack of images. When they found the artifact, they had to stop and notify the instructor. After two test stacks the participants completed the System Usability Scale (SUS) and meCUE questionnaires and started with the next interaction method. The presentation order of the interaction methods was randomized. After finishing all three interaction methods the participants completed a demographic questionnaire.

**Results and Discussion**

Participants were recruited from local hospitals. Overall, 20 surgeons volunteered to participate in this research. From this sample, 10 surgeons participated in a pre-study and the iterative user-centered development process and were therefore excluded from further evaluations. That left 10 surgeons (9 male, 1 female) from different hospitals to take part in the final study. The age ranged from 28 to 55 years (M 39.5, SD 9.62). The average work experience ranged from 4 to 27 years (M 13.1, SD 9.36). 9 of 10 participants were right handed.

<table>
<thead>
<tr>
<th>Sub-scale</th>
<th>Technique</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability</td>
<td>Foot</td>
<td>5.83</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>Hand</td>
<td>6.09</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>UpDown</td>
<td>6.23</td>
<td>0.51</td>
</tr>
<tr>
<td>Usage Intention</td>
<td>Foot</td>
<td>4.06</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Hand</td>
<td>3.97</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>UpDown</td>
<td>4.47</td>
<td>0.69</td>
</tr>
<tr>
<td>Usefulness</td>
<td>Hand</td>
<td>4.96</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>UpDown</td>
<td>5.70</td>
<td>1.09</td>
</tr>
<tr>
<td>Overall Opinion</td>
<td>Foot</td>
<td>2.35</td>
<td>2.04</td>
</tr>
<tr>
<td></td>
<td>Hand</td>
<td>2.95</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>UpDown</td>
<td>2.95</td>
<td>1.36</td>
</tr>
</tbody>
</table>

Table 1: Means and standard deviations for meCUE. Sub-scales range from 1 to 7 while the overall opinion ranges from -5 to 5.

All participants successfully carried out the presented tasks for all interface conditions. When asked about their personal preference for one of the presented techniques, four participants favored foot interaction, three participants favored hand interaction, and three participants favored verbal commands (UpDown). The results for SUS are presented in Table 2.

Comparing the means using a one-way repeated measures ANOVA showed no significant differences across the three techniques, although we found a marginally significant linear contrast (p = 0.056). In absolute numbers, the verbal control technique (UpDown) was rated best, foot interaction was rated worst and hand interaction in between the other

---

**Apparatus**

A list of the used hardware and software

- Two screens
- HTC VIVE
- Leap Motion Controller
- Optitrack Prime 13 system with 2 cameras
- Pair of modified OR-shoes
- Computer running VR environment with Intel Core i7-4790, 16 Gigabytes of RAM, and an NVIDIA GeForce GTX980
- Computer for foot tracking Intel core i7-3770, 16 Gigabytes of RAM, and an NVIDIA GeForce GTX680
- Unreal Engine of Epic Games
- “VrpnPlugin” of the Hochschule Reutlingen
two techniques. However, some participants noted that the verbal commands worked better in the test setting than in the real life scenario.

The results for the meCUE scales are presented in Table 1. Again, comparing the means using a one-way repeated measures ANOVA showed no significant differences. In absolute numbers, the verbal control technique (UpDown) achieved slightly higher scores across all sub-scales than foot and hand interaction.

The results for the average task completion times are presented in Table 3. The measured times of one participant were removed from the data set before calculating the means and excluded for all further analysis as they were more than two standard deviations above the average. Comparing the means using a one-way repeated measures ANOVA showed no significant differences. In absolute numbers, participants achieved the fastest average time using the verbal control technique (UpDown) and a slightly slower time using hand interaction, while the time achieved using foot interacting was the slowest of the three.

While the verbal control technique performed best in absolute numbers, no significant differences could be found across the conditions. Especially, considering the limited acclimatization time, it is noteworthy that all participants were able to use all techniques successfully in all cases and that all techniques were rated very favorably for their usability and user experience. In general, surgeons commented positively on the VR system and the presented interaction techniques.

**Conclusion and Future Work**
In this paper, we presented IVOR and how it can be employed to conduct user-centered design, development, and research for novel interaction techniques inside the OR.

Additionally, we presented and discussed the interaction design of two post-WIMP interaction techniques for browsing medical images inside the OR that were developed and evaluated using IVOR.

We argued how IVOR simplifies the logistics of conducting user studies within a situated context and facilitates iterative development and gathering empirical data on new interaction approaches. Our results show that the IVOR was well received by surgeons who had no prior experience with VR technology.

IVOR was used to develop and evaluate post-WIMP interaction techniques against verbally instructing a human assistant as is currently established practice in many hospitals. The results for the presented foot and hand interaction gestures show that all techniques were used successfully by the participants with only minimal time for acclimatization and no statistically significant differences regarding usability or task completion time.

In the future, we plan to further validate our approach by comparing the results to evaluations of the same techniques outside IVOR and we want to further refine the post-WIMP interaction techniques and integrate them with existing OR devices.

**Acknowledgments**
We would like to thank the hospitals Klinikum Bremen-Mitte, Puis Hospital Oldenburg, Klinikum Bremen-Ost, Rotes Kreuz Krankenhaus, Asklepios Klinik Barmbek and the participants that volunteered to take part in the user study or pre-tests. The work was partly funded by the DFG within the context of the institutional strategy of the University of Bremen in the context of the German "Exzellenzinitiative des Bundes und der Länder".
REFERENCES


