An Immersive System for Exploring and Measuring Medical Image Data

P. Saalfeld, J. Patzschke, B. Preim

To cite this version:

Saalfeld, P., Patzschke, J., Preim, B. (2017), An Immersive System for Exploring and Measuring Medical Image Data. Mensch und Computer, in print

An Immersive System for Exploring and Measuring Medical Image Data

Patrick Saalfeld[1], Johannes Patzschke[1], Bernhard Preim[1]

Visualization Group, Otto-von-Guericke University Magdeburg, Germany[1]

Abstract

We present an immersive system to explore and measure medical 2D and 3D data for treatment planning. Our focus lies on interaction techniques, which are oriented on the workflow of treatment planning. We use the head-mounted display from an Oculus Rift DK2 as output device. For input, we use a stylus from a zSpace and implement ray-based interaction techniques. Our system is designed to be used in a sitting position at a desk, similar to the treatment planning environment of a physician. The stylus is grabbed equally to a pen and allows precise input, which is important for measurements. We evaluated our prototype with computer scientists, stating exploratory and measurement tasks on medical 2D and 3D data. Here, we assessed the time, usability, felt presence and informal feedback. Our results show that our exploration and measurement techniques are fast, easily understood and support presence. However, tracking losses due to infrared interference have to be solved for further studies with physicians.

1 Introduction

Medical diagnosis and intervention planning is mostly done with 2D image data (Reitinger et al., 2006). This data is the foundation for physicians to mentally build 3D models of different structures. Additionally, there exists a wide variety of 3D visualizations of medical structures, which are generally inspected on 2D displays. These displays cannot convey important depth cues such as motion and binocular parallax and must be substituted with other depth cues, e.g., shading and shadows (Preim, Baer, et al., 2016). An immersive environment realized with a head mounted display (HMD) allows to convey these missing depth cues and can present 3D structures more naturally. Additionally, a precise input method is necessary for medical treatment planning tasks such as measurement and exploration. Traditional input devices such as a mouse would allow precise input, but limit the degrees of freedom (DoF) and are not practically usable with HMDs. Modern input technologies such as hand tracking are more natural and can lower the cognitive distance between user input and system feedback. However, tracking inaccuracies and user-generated noise lead to less precise input (Bowman et al., 2004).

This work presents interaction techniques to explore medical 2D and 3D structures and allow distance and angle measurements. These techniques are realized in an immersive environment with the HMD *Oculus Rift* to enhance the spatial understanding. For precise input, a stylus with 6DoF from a *zSpace* is used. This pen-like device makes use of the so-called precision grip, which allows a high accuracy for input (Napier, 1956). The limited tracking area of the zSpace is acceptable, since the application is realized according to a physician's typical treatment planning environment, i.e., sitting on a desk in front of a display.

We evaluated our system by giving 13 participants various exploration and measurement tasks on medical structures in 2D and 3D. We measured the necessary time and collected data of a usability and presence questionnaire. Additionally, the think-aloud method was used to get informal feedback. Our results show that the interaction techniques allow a fast exploration and measurement that is easily understandable and supports presence.

2 Medical Background & Related Work

This section comprises information regarding medical image acquisition, immersive systems for treatment planning as well as related research regarding computer-assisted interactive measurements. Different image acquisition modalities, such as computer tomography (CT) or magnetic resonance imaging (MRI), produce a series of 2D images. By combining a stack of 2D images, patient-specific 3D models can be created. In medical routine, both 2D images as well as the 3D reconstructions are important. Although 3D structures convey the overall shape, size and relations to other structures better, the inside is occluded (Preim and Botha, 2013). 2D images overcome this limitation by revealing the inside and, thus, more details.

For diagnostics and treatment planning, physicians rely on these images to make informed decisions. Furthermore, they must be able to interact with the acquired data in an efficient manner. For medicine, immersive systems are especially valuable, since they allow a realistic representation of 3D data and the usage of natural depth cues such as motion and binocular parallax. Measurements play a key role in diagnostics and treatment (Preim and Botha, 2013). Safety margins between tumors and vessels are important to assess the feasibility of interventions or the risk of surgery. Preim, Tietjen, et al. (2002) investigated different methods to measure distances, angles and volumes on medical data and how to visualize them. Furthermore, purely interactive and fully automatic methods are discussed. However, the presented methods are implemented for 2D input and output devices. In contrast, Reitinger et al. (2006) presented the *LiverPlanner*, a virtual reality system to support physicians in planning liver resections. Their system uses a large stereoscopic projector in combination with tracked shutterglasses and input devices. They present tools to measure distances and angles by defining two and three spatial points, respectively. However, the points are defined without snapping support, which gives more freedom but bears the risk to be less accurate. Furthermore, they present a jug tool that allows measuring volumes by dragging 3D objects into the physical, tracked jug.

3 Exploration and Measurement of Medical Data

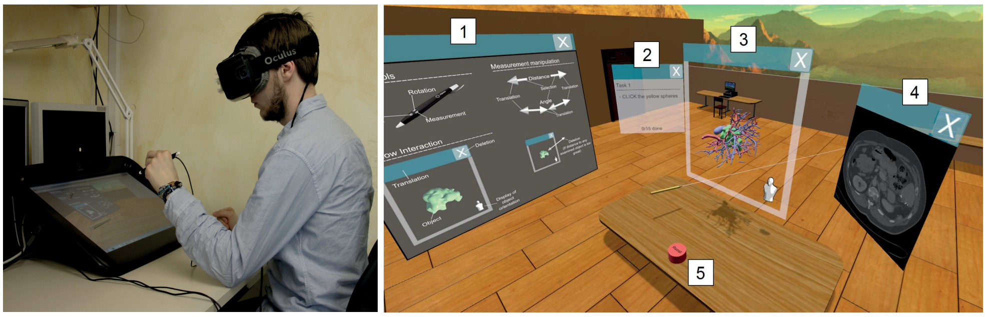
This section discusses design decisions and technical aspects of our system. Furthermore, we describe in more detail how medical 2D and 3D data is presented, explored and measured.

3.1 Design Decisions and Technical Aspects

Immersive Environment. By providing physicians with a life-like representation of 3D structures and natural interaction techniques, their efficiency and understanding can be enhanced. An immersive environment can provide a natural representation by using additional depth cues. Through this spatial perception, reconstructed anatomical models can be visualized, inspected and explored in a more intuitive way (Gallo et al., 2008). We use the HMD *Oculus Rift DK2* to create this environment (see Fig. 1). The HMD's display has a resolution of 960 × 1080 pixels per eye and is tracked with 6DoF. The tracking is realized with a combination of inertial sensors (gyroscope, accelerometer, magnetometer) and an external camera that detects infrared light emitted by LEDs on the HMD. A possible alternative would be the stereoscopic display of a zSpace. Since it supports binocular and motion parallax, it would also enhance the spatial perception. However, this fish tank VR (Ware et al., 1993) device is less immersive, since it restricts the additional depth cues to the area of the stationary display.

Precise 3D Input. The possible precision of an input device is restricted by the way it is held. A general subdivision of gripping techniques is the distinction of *power grips* and *precision grips*. As the name suggests, the power grip, used for tools such as a hammer, is less precise, since the movement is controlled via shoulder and arms (Napier, 1956). Therefore, we decided to use a device that makes use of the precision grip, i.e., the stylus of the zSpace (see Fig. 1). This pen-like device, held with the tripod grip, allows the precision normally used for writing and drawing (Song et al., 2011). Another benefit is its low weight (40 g). During long therapy planning sessions, a heavy input device would probably be rejected by clinicians (Patel et al., 2006). The stylus is tracked with 6DoF, combining active optical tracking with two infrared LEDs providing 5DoF. The sixth DoF is provided by an integrated gyroscope. The tracking resolution is documented with 2 mm (x/y/z) and 2° (yaw/pitch/roll). The positional accuracy is specified with 3/3/5 mm for x/y/z and rotational accuracy with 2° for yaw/pitch/roll, respectively. However, the tracking space is limited by the placement of the zSpace. Usually, this range is implicitly communicated to the user, since the interaction happens in front of the zSpace. In our case, the user wears an HMD. Thus, the tracking space has to be visualized inside the virtual environment. Therefore, we show a warning and semi-transparent wall if the user moves the stylus to the borders of the tracking area. Another possibility would be using input devices with a larger tracking area, e.g., the HTC Vive controller. Although their accuracy in tracking is comparable, they are less precise in handling due to the power grip. Additionally, the larger weight of 203 g would lead to faster fatigue of the user's hand.

Desk Environment. Since clinical routine is connected with high time pressure, the time to familiarize with a new system is critical. Therefore, our virtual environment is oriented on the treatment planning environment of a physician, i.e., sitting at a desk in front of a screen. The seating position is less tedious in longer planning sessions. The virtual desk has the same height



(a)

(b)

Figure 1: Our immersive system to explore and measure medical data in use: the Oculus HMD is used as output device. For precise input, a tracked stylus from zSpace is used that enables mid-air interaction (a). The virtual environment (b) comprises an overview of the functionality (1), textual notes (showing instructions during the evaluation) (2), representation of 3D structures (3), radiological 2D image data (4) and a desk with a reset button (5). The virtual desk has the same height as the real one.

as the real desk, enabling the physician to rest the elbows on it intuitively, without putting down the HMD (see Fig. 1). The resting elbows in combination with the precision grip facilitate convenient and precise measuring with the stylus (Reitinger et al., 2006). By positioning medical data in arms reach, proprioception, i.e., the sense of position and orientation of the user's body, leverages the precision and sense of depth (Mine et al., 1997).

3.2 Used Frameworks

Our system is developed with the game engine Unity (Unity Technologies, San Francisco, USA). This allows us to improve depth perception and felt presence by using rendering features such as real-time global illumination and physically-based rendering. Furthermore, existing frameworks can be used for the Oculus Rift and zSpace integration. However, there is no built-in functionality to load and display medical 2D image data. Therefore, we used the open source C# library *Evil DICOM*. The slices of a data set are read pixel-wise and stored as textures. These textures are mapped onto planes to visualize the medical data.

3.3 Views

We organize our 3D structures, 2D image slices and user information in different dedicated views, which are visualized as borders around the content (see Fig. 1). The interaction with these views follows WIMP concepts (windows, icons, menu, pointer). Since the WIMP concept is still a familiar component of current applications, users' knowledge of interaction methods can be used immediately for our views (van Dam, 1997). Views can be picked by the stylus on the title bar and translated to other locations. The border is always facing the user to prevent perspective distortions on the title bar text or icons. An "X" button on the right of the title bar known from windows on desktop applications allows the user to hide the view.

3.4 Selection

For selection with the stylus, we investigated different techniques, i.e., grabbing and arm extension methods for remote objects. Since the zSpace is positioned in front of the user, the arm cannot be fully extended without hitting the zSpace or leaving the stylus tracking area. Therefore, we decided for a ray-based interaction method. The stylus is represented by a pencil that is shooting a semi-transparent ray into infinity. Even small, remote objects can be selected, since they are contained in a view with an ensured minimum size. If a view or its content gets hit by the pencil or the ray, the view is highlighted and the stylus vibrates to give visual and vibro-tactile feedback.

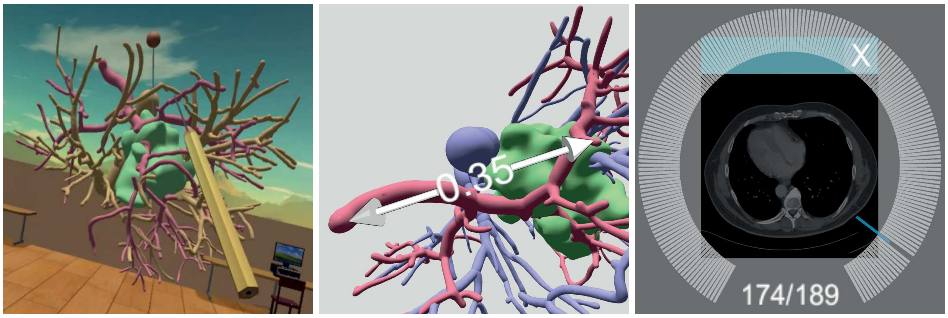
3.5 Translation & Rotation

For more control to transform an object, we divided translation and rotation into two interaction methods. For translation of views and objects, we implemented the *Mesh-Grab* method (Katzakis et al., 2013). After hitting a structure with the pencil or ray and pressing a dedicated stylus button, the movement offset of the hitting point H_P is transferred to the object.

In contrast to translation, rotation techniques are less intuitive. For object rotation, an easy to learn and natural technique would be an egocentric method. Here, the rotational changes of the input device are directly applied to the object rotation. However, cables of input devices and anisotropic sampling make an equally precise rotation in every direction difficult. For example, the rotation of the zSpace stylus around its longitudinal axis is difficult. Therefore, we used an improved version of *Mesh Grab* (Katzakis et al., 2013). After a dedicated button on the stylus is pressed, the hitting point H_P between the mesh and the ray is stored. The mesh center M_C and H_P define the vector $\mathbf{v} = \overrightarrow{M_C H_P}$. After moving the stylus with the pressed button, the vector $\mathbf{v}$ is updated as $\mathbf{v_{New}}$. The rotation axis $\mathbf{r}$ is extracted as $\mathbf{r} = \mathbf{v} \times \mathbf{v_{New}}$ and the rotation angle α is defined as $\alpha = \angle(\mathbf{v},\ \mathbf{v_{New}})$. We improved this technique with two additions. First, if the ray leaves the mesh, the length of $\mathbf{v}$ is increased, resulting in an invisible sphere with increasing radius. Second, we improved the visualization of the technique by drawing the vector $\mathbf{v}$ with a line and representing H_P with a red sphere (see Figure 2). These adaptations have several benefits. They allow different degrees of rotation precision depending on the distance of the ray to the object. Larger distances allow a more precise rotation, since the resulting changes are smaller on a bigger surface of the invisible sphere. Furthermore, objects can be rotated around 180° in a fast way by piercing through the object's center. The visualizations help to understand the rotation by representing the invisible sphere's size and the center of rotation.

3.6 Measurement on 3D Models

As mentioned in Section 2, a series of 2D image slices can be combined to create patient specific 3D models. These objects convey the shape and relations to other objects spatially, which is enhanced due to the immersive environment. We reconstruct our 3D models in a preceded step out of the 2D image data. Distance and angle measurements are created by hitting the object with the ray. After pressing the third stylus button, a distance measurement is created. With the rubber banding interaction technique, a measurement arrow can be dragged to an arbitrary



(a) (b) (c)

Figure 2: Rotation is realized with the Arcball 3D method (Katzakis et al., 2013) (a). The visualization is improved by drawing a line between the rotation center and the gripping point of the object. (b) depicts the cel-shading used for our measurement visualization. In (c), the representation of medical 2D data is shown. The radial interface allows fast slicing through the dataset.

location. To allow visibility even for small measures, the arrows adapt their size according to their length. The measured value is visualized during dragging. Depending on the arrow size, the value is either positioned over the arrow or between the arrow heads. Angle measurements are created based on distance measurements by *grabbing* an arrow in the center. Now, the user drags the third point, i.e., vertex point of the angle. Therefore, no additional button is necessary to differentiate between angle and distant measurements.

All points of a measurement can be adjusted by grabbing them with the ray. Besides adjustment, this also allows to delete a measurement. If the user drags a measurement further away than a specified threshold, a message is shown, indicating that it is deleted after releasing the stylus button. To allow a clear differentiation between medical structures and measurement, the arrows are rendered with a cel-shading (see Figure 2).

3.7 2D Medical Image Data

The image data is shown in a view such as the 3D objects. In contrast to 3D objects, the 2D views cannot be rotated, which is not necessary, since they are always facing the user. For each anatomical axis (sagittal, coronal, transverse) a separate view is available. To change the currently visible slice, we implemented a technique inspired by marking menus (Kurtenbach and Buxton, 1994). After pressing a dedicated button, a ring is shown around the 2D view. Every slice is visualized as a ring segment (see Figure 2). By pointing with the ray on one of these segments, the corresponding slice is shown. This representation allows fast jumps into different areas of the data set, which is not possible with the common incremental slicing via the mouse wheel.

4 Evaluation

For our evaluation, physicians would be ideal participants, since they represent the target group. However, no clinicians were available for the study, thus, we evaluated our immersive system with computer scientists with experience in medical applications. To obtain results that are adaptable to clinicians, we used a medical data set of a liver, comprising supplying vascular structures and a tumor. However, our designed exploration and measurement tasks were solvable without medical background knowledge. During our evaluation, we collected:

- parametric data in form of time during exploration and measurement tasks,
- non-parametric data with questionnaires regarding usability, felt presence and experience in related topics and
- informal feedback with the think-aloud method.

For usability, we used a questionnaire from Prümper (1997) that is oriented on the ISO 9241-110 standard. For felt presence, we used the *control factors* from the questionnaire of Witmer and Singer (1998). All questions were stated with a 7-point Likert scale from -3 to 3.

4.1 Procedure & Tasks

At first, all participants were introduced to the input and output devices as well as to the immersive environment. They could take as long as they want to make themselves comfortable with the immersive environment. The participants had to solve three types of tasks that became more difficult successively:

1. Yellow spheres appeared one by one at predefined positions and slices, respectively. The task was to find these and select them. Here, translation and rotation of objects or changing the slice of the image data set is necessary, which represents the exploratory aspect.
2. Groups of two yellow spheres appeared at predefined positions or slices. The participants had to measure the closest possible distance between these spheres.
3. Three spheres appeared, two in yellow, one in red. The participants had to create angle measurements by first creating a distance measurement between the two yellow spheres and then transforming it to an angle measurement by dragging it to the red sphere.

To enable the participants to explore and search inside a stack of image slices without inspecting every slice, we highlighted an area of possible segments in the ring visualization. This highlighting recreates the physicians' knowledge to a certain extent, since they know the approximate region to investigate inside the image stack. No handout was possible; therefore, one dedicated view showed the current task in the immersive environment. The participants were encouraged to share their thoughts on problems, surprising elements and favored features. After finishing these tasks, the participants answered the questionnaires.

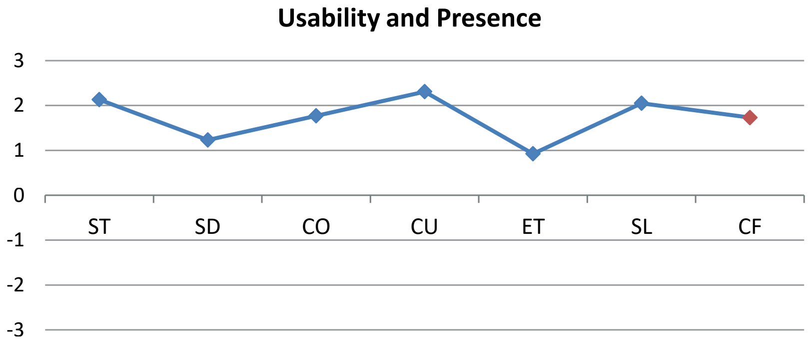


Figure 3: Overview of the results of the usability and presence questionnaire (ST – suitability for the task, SD – self descriptiveness, CO – controllability, CU – conformity with user expectations, ET – error tolerance, SL – suitability for learning, CF – control factors).

4.2 Results

Overall, 13 computer scientists (3 women, 10 men) took part in our evaluation. On average, they were 29.7 years old (standard deviation $s = 6.6$), had 18.1 ($s = 6.0$) years of experience with computers, 5.2 ($s = 6.4$) years with medical data, 2.7 ($s = 4.8$) years with 3D UI interaction and 1.5 ($s = 1.8$) years with virtual reality.

Time. Regarding the tasks on the 3D objects, the participants required $\bar{x} = 6\,\mathrm{s}$ ($s = 2.9$) for the selection task, $\bar{x} = 36.4\,\mathrm{s}$ ($s = 3.5$) for distance measurements and $\bar{x} = 43.9\,\mathrm{s}$ ($s = 2.0$) for angle measurements. For the 2D slices, selecting yellow spheres took $\bar{x} = 14.4\,\mathrm{s}$ ($s = 6.7$), measuring distances $\bar{x} = 26.7\,\mathrm{s}$ ($s = 2.3$) and measuring angles $\bar{x} = 29.6\,\mathrm{s}$ ($s = 2.3$). Interestingly, the measurement tasks took longer on the 3D objects but not the exploration task. A possible explanation for this is that 3D objects can be explored more naturally in the immersive environment. A small head movement can be enough to find a yellow sphere. On the other hand, the 2D slices had to be investigated one by one in the highlighted region.

Usability & Presence. The results of the usability questionnaire were overall positive (see Figure 3). The usability was rated in average with $\bar{x} = 1.7$ ($s = 0.5$). The areas *suitability for the task* ($\bar{x} = 2.1$, $s = 0.2$) and *conformity with user expectations* ($\bar{x} = 2.3$, $s = 0.4$) had the best results. In contrast, *self-descriptiveness* ($\bar{x} = 1.2$, $s = 0.3$) and error tolerance ($\bar{x} = 0.9$, $s = 1.0$) performed worst. One possible reason for these results could be that the information view on how to use the system was rarely used by the participants. Another likely possibility for these ratings are periodically occurring problems with the stylus tracking. After the stylus was lost by the zSpace, it could take several seconds until the participants could proceed using it, which led to frustration.

Regarding the presence questionnaire, the category *control factors* represents the perceived degree of control interacting in the immersive environment, which supports presence. Since our participants rated this positively with $\bar{x} = 1.7$ ($s = 0.6$), it can be assumed that our participants felt immersed.

Think Aloud. The interaction techniques were commented positively overall. The stylus representation in the immersive environments was perceived as very realistic. Additionally, the participants liked the precision and control that was possible with the stylus. However, the periodical tracking losses of the stylus spoiled this impression.

The creation of measurements by placing arrows was commented as easy and intuitive. Some participants suggested improving the visualization of the angle measurement by visualizing an arc between the sides of the angle. In addition, if the distance measurement was too small, participants had problems to create an angle measurement out of it.

The selection of a slice in the 2D image data view was commented as novel and intuitive. Although the ring interface allows a fast selection of a general slicing area, changing single slices was more difficult due to hand tremble. The participants suggested adding a possibility to slice through the data set in single steps.

5 Summary

The presented immersive system allows an interactive exploration and measurement of medical 2D and 3D data. We used ray-based interaction techniques and took inspiration from the WIMP paradigm to add a familiar component for the users.

Our evaluation shows that our system was perceived usable, intuitive and supports presence. In general, our tool has to be evaluated with clinicians to identify extensions for treatment planning. Here, measurement accuracy has to be taken into account. Deviations of a few degrees or millimeters have different meanings in distinct medical domains. Thus, it is important to evaluate our system in a concrete medical domain.

From a technical point of view, the biggest problem was the tracking losses of the zSpace's stylus. These are most likely caused by interferences between the infrared tracking of the Oculus and zSpace. A possible solution is to use another input device or another HMD. Since the precision grip of the stylus allows more accurate input, another HMD is the favorable solution.

For the 2D view, a possibility to change single slices, e.g., realized over additional buttons on the ring interface, would improve the selection. Further extensions would be to allow annotations on the 2D and 3D data. The stylus would be predestined for handwritten and sketched annotations. Additionally, a registration between 2D and 3D data would be beneficial, since it would allow an easier finding of corresponding structures.

References

Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). 3D User Interfaces: Theory and Practice. Addison Wesley Longman Publishing Co., Inc.

Gallo, L., De Pietro, G., & Marra, I. (2008). 3D Interaction with Volumetric Medical Data: Experiencing the Wiimote. In Proc. of Ambient Media and Systems (14:1–14:6).

Katzakis, N., Seki, K., Kiyokawa, K., & Takemura, H. (2013). Mesh-Grab and Arcball-3D: Ray-based 6-DOF Object Manipulation. In Proc. of Computer Human Interaction (pp. 129–136).

Kurtenbach, G. & Buxton, W. (1994). User Learning and Performance with Marking Menus. In Proc. of Human Factors in Computing Systems (pp. 258–264).

Mine, M. R., Brooks, F. P., Jr., & Sequin, C. H. (1997). Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. In Proc. of Computer Graphics and Interactive Techniques (pp. 19–26).

Napier, J. R. (1956). The Prehensile Movements of the Human Hand. Bone & Joint Journal, 38-B(4), 902–913.

Patel, H., Stefani, O., Sharples, S., Hoffmann, H., Karaseitanidis, I., & Amditis, A. (2006). Human Centred Design of 3-D Interaction Devices to Control Virtual Environments. International Journal of Human-Computer Studies, 64(3), 207–220.

Preim, B., Baer, A., Cunningham, D., Isenberg, T., & Ropinski, T. (2016). A Survey of Perceptually Motivated 3D Visualization of Medical Image Data. Computer Graphics Forum, 35(3), 501–525.

Preim, B. & Botha, C. (2013). Visual Computing for Medicine. Morgan Kaufmann Publishers.

Preim, B., Tietjen, C., Spindler, W., & Peitgen, H.-O. (2002). Integration of Measurement Tools in Medical Visualizations. In IEEE Visualization (pp. 21–28).

Prümper, J. (1997). Der Benutzungsfragebogen ISONORM 9241/10: Ergebnisse zur Reliabilität und Validität. In Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung (pp. 253–262). Vieweg+Teubner Verlag.

Reitinger, B., Bornik, A., Beichel, R., & Schmalstieg, D. (2006). Liver Surgery Planning Using Virtual Reality. IEEE Computer Graphics and Applications, 26(6), 36–47.

Song, H., Benko, H., Guimbretiere, F., Izadi, S., Cao, X., & Hinckley, K. (2011). Grips and Gestures on a Multi-touch Pen. In Proc. of Human Factors in Computing Systems (pp. 1323–1332).

van Dam, A. (1997). Post-WIMP User Interfaces. Commun. ACM, 40(2), 63–67.

Ware, C., Arthur, K., & Booth, K. S. (1993). Fish tank virtual reality. In Proc. of human factors in computing systems (pp. 37–42).

Witmer, B. G. & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoper. Virtual Environ. 7(3), 225–240.