Touchless Measurement of Medical Image Data for Interventional Support

P. Saalfeld, D. Kasper, B. Preim, C. Hansen

Touchless Measurement of Medical Image Data for Interventional Support

Patrick Saalfeld[1], Dominique Kasper[2], Bernhard Preim[1], Christian Hansen[2]

Visualization Group, Otto-von-Guericke University Magdeburg, Germany[1]
Computer-Assisted Surgery Group, Otto-von-Guericke University Magdeburg, Germany[2]

Abstract

The preservation of sterility is essential during interventions. Based on interviews with physicians and observed interventions, we derive requirements for touchless distances measurements. We present interaction techniques to apply these measurements on medical 2D image data and 3D planning models using the Leap Motion Controller. A comparative user study with three medical students and eleven non-medical participants was conducted, comparing freehand gesture control with the established, but non-sterile mouse and keyboard control. We assessed the time, accuracy and usability during 2D and 3D distance measurement tasks. The freehand gesture control performed worse compared to mouse and keyboard control. However, we observed a fast learning curve leading to a strong improvement for the gesture control, indicating that longer training times could make this input modality competitive. We discuss whether the advantage of sterility of gesture control can compensate for its inferior performance.

1 Introduction

Evaluation of anatomical and pathological structures in interventional radiology is based on 2D medical image data. Besides a qualitative visual assessment, quantitative measurements are a necessity for treatment decisions (Rössling et al., 2010). For example, tumor treatment depends on the tumor's distance to the surrounding tissue, its volume and its maximum extent. Therefore, the possibility to measure different properties during an intervention is beneficial. To maintain the aseptic environment, the physician's interaction with the image data has to be sterile. In general, there exist two approaches to maintain sterility. First, the physician directly interacts with the image data, using wrapped input devices or scrubbing in again after breaking asepsis. Second, the physician delegates the interaction to a medical assistant with voice and gesture commands. Both approaches can be inaccurate, user-hostile, time-consuming, interrupting and error-prone (O'Hara et al., 2014).

There exist a wide variety of research projects investigating the first approach with touchless input devices. However, they mostly deal with basic interaction tasks of medical image viewers such as rotation and zooming (Mewes et al., 2017). Our goal is to investigate more advanced tasks, i.e., measuring distances within medical image data. We present sterile interaction techniques to create distance measurements with the Leap Motion Controller (LMC). We used the LMC since it was successfully used in interventional settings (Mewes et al., 2017). Furthermore, we use strategies to support the physician to memorize and execute gestures. We evaluated our gestures by comparing them to non-sterile mouse and keyboard interaction, considering aspects such as necessary time, accuracy, usability and tiredness. Our results show that our gesture control was inferior to mouse and keyboard interaction. However, considering that measuring within medical image data can be performed in a usable and sterile manner, our gestures are still beneficial regarding patient safety.

2 Medical Background & Related Work

Medical image acquisition modalities, most significantly computer tomography (CT) and magnetic resonance imaging (MRI), produce a series of 2D images. These image stacks can be combined to create patient-specific 3D planning models. Both 2D images and 3D planning models are necessary in medical routine.

In contrast to open surgery, interventional radiology aims at minimally-invasive procedures. Due to the small incision, imaging control is mandatory to guide the physician. Therefore, several research projects investigate a sterile possibility to interact with medical image data. An adequate way to do this, is the usage of gestural interaction techniques. In general, the accurate and reliable recognition of 3D hand poses and gestures remains a challenging research area (LaViola, 2013). Since passive vision-based sensors allow an unobtrusive tracking of hands, they are a common choice where users are not able to hold a device. Thus, these sensors are used in this work. Gestural interaction is also challenging due to missing haptic feedback and larger space requirement. These and further pragmatic (effective and efficient goal-achievement) and hedonic qualities (fun and aesthetics) differ compared to other interaction techniques. This was shown in the study of van Beurden et al. (2012), where gesture interaction performed significantly worse regarding *perceived performance* and *pragmatic quality*, but better regarding to hedonic qualities. To improve the downsides of gestural interaction, this work supports the user to memorize and execute gestures.

An example for the usage of gestural interaction is the work of Riduwan et al. (2013). They used the Microsoft Kinect tracking as an input modality. They realized basic interaction tasks such as pointing and rotation. However, no evaluation was performed, which is a general problem according to the literature review of Mewes et al. (2017). A comparative user study is presented by Saalfeld et al. (2015). They compared touchless interaction with touch input during basic interaction tasks with a medical image viewer. Their evaluation showed a significantly better performance and intuitiveness for the touch screen interaction.

The described systems allow basic interaction tasks with medical image data; however, measurements were not investigated. One exception is the work of Rosa and Elizondo (2014). They presented freehand gestures for dental surgery procedures. Their LMC-controlled system allowed to control basic medical image viewer functionality and the creation of distance

Figure 1: Exemplary setup of an interventional operation room at the neuroradiological institute at the university hospital of Magdeburg.

measurements. Other work focused on interactive measurement of 2D and 3D structures in the medical domain, however, without using freehand gestures. An example for this is presented by Reitinger et al. (2006). They use a 3D input device for surgical planning that supports to measure distances, angles and volumes.

3 Requirement Analysis

Our requirement analysis is based on two structured interviews (45 minutes each) with physicians and observations of two interventions. Both were conducted at the neuroradiological institute at the university hospital of Magdeburg. Out of this information, we derived requirements in an iterative process together with the interviewed physicians. During the interventions, aneurysms were treated by *coiling*, i.e., a platinum wire was used to prevent blood flowing inside the aneurysm to avoid rupture. Here, vessel diameters, lengths, aneurysm heights and volumes are important measurements to select the correct treatment method. Figure 1 shows the setup of an interventional operation room. Besides non-functional requirements, such as sterility, usability and joy of use, we identified the following functional requirements for measurement tasks.

Simultaneous 2D/3D presentation. Depending on the task, measurements can yield more useful results on the 2D image data or on 3D planning models. For example, the heterogeneity of a tumor has to be inspected on a 2D slice, but the spatial extent can be measured more easily on a 3D model (Preim et al., 2002). Therefore, our system should show the 2D image data in all three standard 2D image directions (sagittal, frontal, transverse) as well as the 3D planning models.

Basic Interaction. During the observed interventions, the physician started by loading the image data and exploring it to find the correct perspective for measurements. During this navigation, rotations were most commonly used. For supporting these functionalities, a possibility to select objects and graphical user interface (GUI) elements is necessary. Furthermore, the physician should be able to translate, scale and slice inside the 2D image data. For the 3D planning models, translation, scaling and rotation should be possible.

Distance Measurement. According to the interviewed physicians, distances are the most important type of measurement. For example, they are used to determine the distance between a tumor and essential risk structures, such as larger vessel. Physicians measure the vertical and horizontal circumference of aneurysms to estimate what kind of coil is appropriate. Therefore, we focus on distance measurements with our system. To cope with possible errors during measurement creation, the physicians should be able to adjust and delete existing measurements.

Precise Positioning of Measurements. The ability to precisely position measurements is crucial for accurate measurements (Preim et al., 2002). This is especially difficult with freehand gestures, since hand tremble and inaccurate tracking creates noise (Hagedorn et al., 2007). Therefore, user support in the form of smoothing and snapping should be available.

4 Touchless Measurement of Medical Image Data

This section presents details of our developed system comprising the technical setup, interaction techniques and approaches to support the physicians to memorize and execute gestures.

4.1 Technical Setup

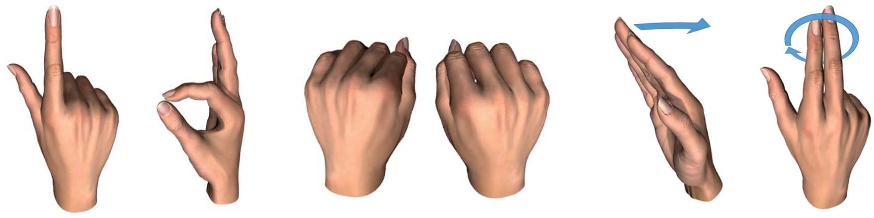
For hand gesture recognition, we used the LMC (Leap Motion Inc., San Francisco, USA). It tracks both hands including single joints with a sampling rate of 39 Hz, a viewing angle of 150° vertically, 120° horizontally and a positional accuracy of 2.5 mm. The used Leap Motion Windows SDK 2.3.1 provides several predefined gestures. Our prototype is developed with the game engine Unity (Unity Technologies, San Francisco, USA). To load the medical image data, the open source C# library *Evil DICOM* was adapted to be used in Unity.

4.2 Interaction Techniques and Gestures

Our gestures can be used with both hands, since bimanual interaction allows more efficient work and an improved perception of the interaction space (Hinckley et al., 1998). Our used gestures are based on previous publications as well as gestures provided by the LMC API.

Pointing. For pointing on different views, objects and buttons, an extended index finger is used (Fig. 2a). By projecting the pointing direction onto the display, ray-based interaction is possible, which was found to be intuitive and minimally tiring for the hand (Fikkert et al., 2010).

Pinch-to-Click. For selecting a GUI element, a structure or to create a distance measurement, a pinch gesture is used (Fig. 2b). This gesture is not executed with the pointing hand, but with the other one. This method was found to be fast, natural and unambiguous (Ni et al., 2011). Furthermore, the kinesthetic feedback triggered by the contact of index finger and thumb generates tactile feedback, which is otherwise missing on freehand gestures.



(a) (b) (c) (d) (e)

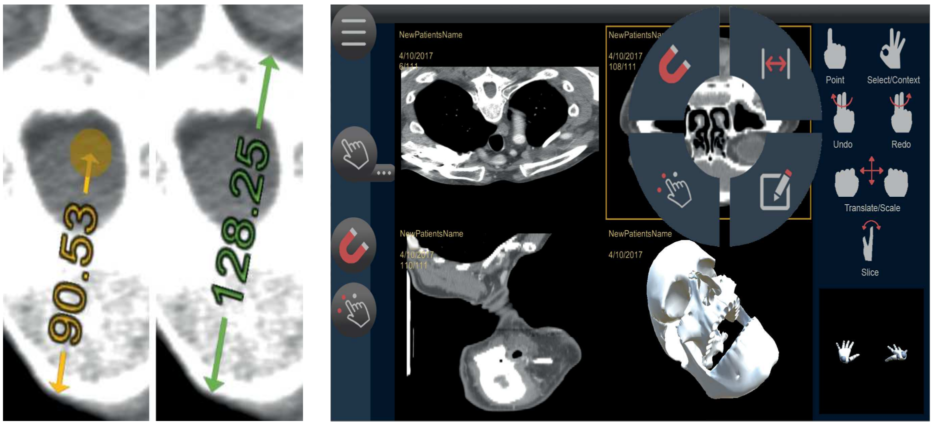
Figure 2: Overview of our used gesture set. To create a measurement, a combination of pointing (a) and placing measurement points (b) is necessary. Object and camera transformations are realized with a handle bar metaphor, where the objects are skewered on a virtual handle bar (c). For scrolling through medical 2D image data, a swipe gesture is used (d). For undo and redo actions, a circle gesture is used (e). All single hand gestures can be executed with either the right or the left hand.

Translation, Rotation and Scaling. All manipulations to transform the camera or objects start with both hands forming the same gesture next to each other. For the camera, both hands form a fist (Fig. 2c) and for objects, the pinch gesture is used with both hands. This mimics the metaphor of objects that are skewered on a bimanual handle bar, which was shown to be precise, efficient and intuitive (Song et al., 2012). Simultaneously moving both hands in any direction translates the camera or object according to the movement. Pitch and yaw rotation is triggered by rotating the handle bar around the corresponding axis. The roll rotation around the handle bar itself cannot be tracked accurately. Therefore, a pedaling motion of both hands is used (Song et al., 2012). To scale objects uniformly, the hands are moved apart or closer together.

2D Medical Image Scrolling. For changing the currently visible image slice, a swipe gesture is used. Here, one hand is vertically held over the LMC and then swiped to the left or right, which shows the previous or next slice (Fig. 2d). To prevent a laborious scrolling of many single images, the gesture can be held at the end of a swipe movement, allowing continuous scrolling.

Creating and Editing Distance Measurements. Distance measurements are created by placing single points with a consecutive usage of the pointing and pinch-to-click gesture. During creation, measurements are colored yellow. After a measurement is placed, it is colored green and can be edited or deleted. First, the physician has to change into *edit mode* by using the radial menu (see Section 4.4). We decided for a mandatory mode change since selecting an existing measurement can be ambiguous regarding other tasks, such as creating a new measurement. After selecting a measurement in edit mode, it is colored in red and its end points are highlighted. Now, with pointing and pinch-to-click, one end point can be repositioned.

Undo and Redo Actions. Besides editing measurements, we cope with errors by allowing the physician to undo and redo different executed actions. For this, we used a circle gesture. First, the physician holds the hand vertically over the LMC with an extended index and middle finger (Fig. 2e). Now, circling the hand clockwise or counter-clockwise triggers the undo and redo of an action, respectively.



(a) (b)

Figure 3: In (a), the measurement process is shown. After placing the first measurement point, the second point is positioned with the rubber band metaphor. If the second measurement position is placed, the finished measurement is colored green. In (b), the overview of our system is depicted. The left bar contains menus to load image data and change measurement settings. In the center, different views on the 2D image data and 3D planning models are available and the radial menu is shown. The right side contains visualizations to support the physicians memorize and execute currently available gestures.

4.3 Support to Memorize and Execute Gestures

Visual Support. We support the physician to learn and memorize the available gestures with three different visual approaches. First, all currently possible gestures are shown with icons and informative names of their functionality on the right side of our application GUI (Fig. 3b). Second, the recognized 3D hand models are visualized at the right bottom. This allows the physician to evaluate if a misrecognized gesture is caused by the LMC or by the execution. Third, the icon representing the pointing position changes according to the recognized gesture. Here, a semitransparent yellow icon is used that does not occlude content.

Algorithmic Support. Besides the visual approaches, algorithmic support is necessary to cope with the imprecise tracking of the LMC. To smooth the tracked hand positions, we use *exponential smoothing*. Here, a specific count n of previously tracked positions y_{t-i} and the current positions y_t are weighted and added, where the weight decreases with older positions. The result is a prediction of the next value, calculated with $y_{t+1} = \sum_{i=0}^{n} \alpha(1-\alpha)^i y_{t-i}$. The parameter α is a weighting factor, resulting in faster reactions and less smoothing with higher values. We empirically determined these values depending on the used gestures.

Additionally, the physician can activate snapping (Hagedorn et al., 2007), i.e., the pointing position snaps to nearby relevant structures. To find snapping points in the 2D image, we used the *Sobel operator*, an image processing filter that extracts edges. Every pixel that belongs to an edge is a possible 2D snapping position. For the 3D data, we use the vertices of the mesh as possible 3D snapping positions. During pointing, a nearest neighbor search is done inside a quadtree (edges) and octree (vertices), respectively, to allow for real-time performance.

Figure 4: The evaluation setup for the gesture pass (left) and for the mouse and keyboard pass (right).

4.4 Graphical User Interface

The user interface of our system is divided into three parts: a sidebar on the left contains system options, the 2D and 3D visualizations of the medical data are positioned in the center and on the right, information about the available and recognized gestures are shown (see Figure 3b).

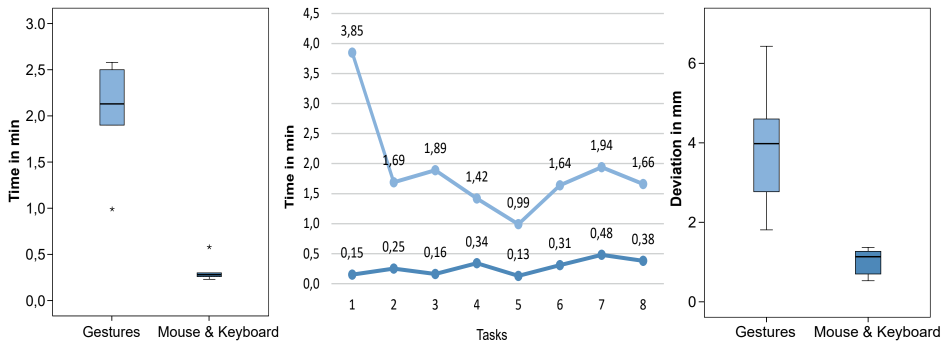
Besides these components, we implemented a radial menu that is used as a context menu (Fig. 3b). Radial menus allow fast and efficient access to hidden functionality without occupying space permanently. Furthermore, it is centered on the current pointing position. Thus, all options can be reached within the same distance (Chertoff et al., 2009). The radial menu is showing up after holding the pinch-to-click gesture. While holding, the hand is moved to the desired function inside the menu. The function is then triggered by stopping the gesture.

5 Evaluation

To compare our gesture set to an established interaction method, we implemented the possibility to control our system with a mouse and keyboard. We collected different parametric data (time, accuracy) as well as non-parametric data (usability, tiredness of the hands). For usability, we used the *System Usability Scale* (SUS) (Brooke, 1996). Additionally, we assessed the tiredness of fingers, wrists, arms and shoulders with questions for each body part. The study was realized on a Sectra table (Sectra AB, Linköping, Sweden) with a 55″ display, which resembles the display available in a radiological intervention room. Depending on the input modality, either the LMC or the mouse and keyboard were placed in front of the participants (Fig. 4).

Participants. Overall, 14 participants took part in our study (8 women, 6 men). On average, they were 25.7 years old. Although we reached out for participants with medical background, only three medical students took part. The other participants were students from mixed domains, including computer science and engineering. However, the recreation of existing measurements was understandable without a medical background, thus, valid results were still obtainable.

Procedure and Tasks. After training, where participants could practice every gesture one by one as long as they want, they had to recreate six predefined distance measurements. These were uniformly distributed on the 2D and 3D views. Furthermore, they were placed on various



(a) (b) (c)

*Figure 5: Overview of the results comparing gesture input with mouse and keyboard regarding the necessary measurement time (a), development of time by consecutive tasks (b) and deviation of measurements (c). The whiskers show the interquartile range (IQR)*1.5.*

positions (directly on borders, slightly beside them or far away from them) to alter the usefulness of the snapping feature. Finally, the participants edited existing measurements. After solving all tasks, the participants answered the questionnaire.

Experimental Design. The participants should perform several measurements on 2D and 3D data. Our test showed that these measurements took about one hour for one input modality. We assessed that one hour would result in still acceptable signs of fatigue but is also an upper limit. Therefore, we have chosen a between-subject design, i.e., each participant only uses either gesture control or mouse and keyboard input. We alternatingly assigned the participants to either of the two groups, resulting in seven participants (four women, three men) in each group.

5.1 Results

Time. After removing one outlier, we compared the times that were necessary to create and edit measurements. Overall, participants took 5.9 times longer with the LMC (see Figure 5a). Interestingly, there is a strong decrease of times observable for gestures (see Figure 5b). This indicates that frequent usage of gesture control with longer training times can reduce the difference to mouse and keyboard control.

Accuracy. After removing one outlier (not the same as for the time), the comparison of accuracy shows that participants measured 3.82 times less accurate with the LMC compared to mouse and keyboard control, which results in a deviation of 3 mm (see Figure 5c). Investigating the accuracy regarding 2D image data and 3D planning models, the accuracy is less precise for both input modalities on 3D data. This is presumably caused by the additional dimension, which makes precise measurements more difficult.

Usability. For the SUS questionnaire, we calculated the overall usability that lies between 0 and 100 (Brooke, 1996). The gesture input resulted in a score of 51.8 and mouse and keyboard control in a score of 68.9, respectively. Both values can be interpreted as *ok* according to Bangor

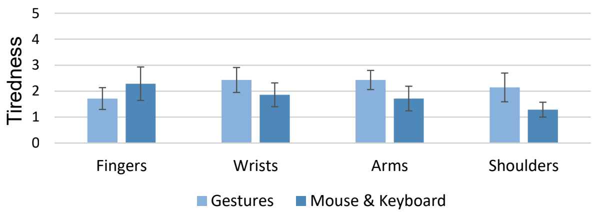


Figure 6: Felt tiredness for fingers, wrist, arms and shoulder with standard error.

et al. (2009). Regarding the tiredness, both input modalities lead to similar results (see Figure 6). The mouse and keyboard control was perceived more strenuous for the fingers, whereas gesture control led to higher tiredness for the wrists, arms and shoulders.

6 Conclusion

We presented interaction techniques to create an important type of measurements for interventional radiology, i.e., distances on medical 2D image data and 3D planning models. Our system fulfills the requirements obtained by interviews with physicians and observations of radiological interventions. The user study showed the inferiority of gestural control compared to mouse and keyboard interaction. The main reasons for this are problems with gesture recognition and an unacquainted input method for our participants. However, our gestures were rated usable according to the SUS scale and the participants were able to create measurements. Furthermore, the time that participants required for measurement creation shortened considerably after the first tasks, indicating that longer training times could improve their performance.

To use our system in a clinical environment, the accuracy has to be improved. For the required time, on the other hand, the gesture control does not necessarily have to compete with mouse and keyboard input. Given that the interaction is sterile and, thus, ensures asepsis, the risk for infection on a patient is strongly reduced. According to statements of our interviewed physicians, higher patient safety can justify longer treatment times.

Although three medical students participated in our evaluation, they do not represent experienced physicians. Therefore, a generalization of our results to a realistic clinical setting is not possible. Thus, an evaluation with physicians in a realistic clinical setting is still necessary. This would also allow to investigate, if additional functionality is necessary.

Acknowledgments. This work is partially funded by the Federal Ministry of Education and Research (BMBF) within the STIMULATE research campus (grant number 13GW0095A).

References

Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. Usability Studies, 4(3), 114–123.

Brooke, J. (1996). SUS: A Quick and Dirty Usability Scale. In Usability Evaluation in Industry. Taylor and Francis.

Chertoff, D. B., Byers, R. W., & LaViola, J. J., Jr. (2009). An Exploration of Menu Techniques Using a 3D Game Input Device. In Proc. of Foundations of Digital Games (pp. 256–262).

Fikkert, W., van der Vet, P., & Nijholt, A. (2010). Gestures in an Intelligent User Interface. In Multimedia Interaction and Intelligent User Interfaces: Principles, Methods and Applications (pp. 215–242). Springer London.

Hagedorn, J. G., Dunkers, J. P., Satterfield, S. G., Peskin, A. P., Kelso, J. T., & Terrill, J. E. (2007). Measurement Tools for the Immersive Visualization Environment: Steps Toward the Virtual Laboratory. Journal of Research of the National Institute of Standards and Technology, 112(5), 257–270.

Hinckley, K., Pausch, R., Proffitt, D., & Kassell, N. F. (1998). Two-handed Virtual Manipulation. ACM Trans. Comput.-Hum. Interact. 5(3), 260–302.

LaViola, J. J. (2013). 3D Gestural Interaction: The State of the Field. ISRN Artificial Intelligence, 1–18.

Mewes, A., Hensen, B., Wacker, F., & Hansen, C. (2017). Touchless Interaction with Software in Interventional Radiology and Surgery: A Systematic Literature Review. International Journal of Computer Assisted Radiology and Surgery, 12(2), 291–305.

Ni, T., Bowman, D. A., North, C., & McMahan, R. P. (2011). Design and Evaluation of Freehand Menu Selection Interfaces Using Tilt and Pinch Gestures. Int. J. Hum.-Comput. Stud. 69(9), 551–562.

O'Hara, K., Gonzalez, G., Sellen, A., Penney, G., Varnavas, A., Mentis, H., … Carrell, T. (2014). Touchless Interaction in Surgery. Commun. ACM, 57(1), 70–77.

Preim, B., Tietjen, C., Spindler, W., & Peitgen, H. O. (2002). Integration of Measurement Tools in Medical 3D Visualizations. In IEEE Visualization (pp. 21–28).

Reitinger, B., Schmalstieg, D., Bornik, A., & Beichel, R. (2006). Spatial Analysis Tools for Virtual Reality-based Surgical Planning. In 3D User Interfaces (pp. 37–44).

Riduwan, M., Basori, A. H., & Mohamed, F. (2013). Finger-based Gestural Interaction for Exploration of 3D Heart Visualization. Social and Behavioral Sciences, 97, 684–690.

Rosa, G. M. & Elizondo, M. L. (2014). Use of a Gesture User Interface as a Touchless Image Navigation System in Dental Surgery: Case Series Report. Imaging Sci Dent, 44(2), 155–160.

Rössling, I., Cyrus, C., Dornheim, L., Boehm, A., & Preim, B. (2010). Fast and Flexible Distance Measures for Treatment Planning. International Journal of Computer Assisted Radiology and Surgery, 5(6), 633–646.

Saalfeld, P., Mewes, A., Luz, M., Preim, B., & Hansen, C. (2015). Comparative Evaluation of Gesture and Touch Input for Medical Software. In Mensch und Computer (pp. 143–152).

Song, P., Goh, W. B., Hutama, W., Fu, C.-W., & Liu, X. (2012). A Handle Bar Metaphor for Virtual Object Manipulation with Mid-air Interaction. In Proc. of Human Factors in Computing Systems (pp. 1297–1306).

van Beurden, M. H. P. H., Ijsselsteijn, W. A., & de Kort, Y. A. W. (2012). User Experience of Gesture Based Interfaces: A Comparison with Traditional Interaction Methods on Pragmatic and Hedonic Qualities. In Gesture and Sign Language in Human-Computer Interaction and Embodied Communication (pp. 36–47). Springer Berlin Heidelberg.