PREPRINT

Activation Modes for Gesture-based Interaction with a Magic Lens in AR Anatomy Visualisation

Lovis Schwenderling[a, b], Anna Kleinau[a], Wilhelm Herbrich[a], Haswanth Kasireddy[a], Florian Heinrich [a, b] and Christian Hansen[a, b]

[a]Faculty of computer science, Otto-von-Guericke University Magdeburg, Germany;
[b]STIMULATE research campus, Magdeburg, Germany

ARTICLE HISTORY
Compiled September 13, 2022

ABSTRACT
Learning human anatomy is key for health-related education and often requires expensive and time-consuming cadaver dissection courses. Augmented reality (AR) for the representation of spatially registered 3D models can be used as a low-cost and flexible alternative. However, suitable visualisation and interaction approaches are needed to display multilayered anatomy data. This paper features a spherical volumetric AR Magic Lens controlled by mid-air hand gestures to explore the human anatomy on a phantom. Defining how gestures control associated actions is important for intuitive interaction. Therefore, two gesture activation modes were investigated in a user study (n=24). Performing the gestures once to toggle actions showed a higher interaction count since an additional stop gesture was used. Holding the gestures was favoured in the qualitative feedback. Both modes showed similar performance in terms of accuracy and task completion time. Overall, direct gesture manipulation of a magic lens for anatomy visualisation is, thus, recommended.

KEYWORDS
magic lens, augmented reality, gesture interaction

1. Introduction

Learning human anatomy is an important part of medical school and other health-related education. Dissection courses using human cadavers are important for learning anatomical variations and spatial relationships of organs (Ghosh 2017). However, cadavers can be cost and time intensive, could present challenges to students for ethical reasons, or simply might not be readily available (Estai and Bunt 2016). The usage of volumetric three-dimensional (3D) models for anatomy education has proven to support spatial understanding and factual knowledge while increasing user satisfaction and subjective efficiency (Yammine and Violato 2015). Here, augmented reality (AR), can be used to further support the learning of anatomical structures and relationships by enhancing the real environment with virtual content. Using AR, 3D models can be displayed spatially registered in a body or phantom to reduce the cognitive load and improve performance of the students (Buchner et al. 2022). Here, projector-based

Anna Kleinau and Lovis Schwenderling contributed equally as first authors

CONTACT: Christian Hansen. Email: christian.hansen@ovgu.de

spatial AR has the advantage of a larger, less constrained display that is more easily accessible for multiple people than, for example, head-mounted displays.

Various, often occluded structures are relevant for learning anatomy. In general, a lot of information must be presented, which in an AR environment quickly leads to visual cluttering and, thus, to an overload of the user (Pindat et al. 2013). Focus and context visualisations present detailed information for a small, limited area while displaying more superficial information around it to maintain the overall picture (Bjork et al. 1999). Magic lenses are a variant of focus and context visualisations, inspired by a magnifying glass, in which supplementary, hidden or enlarged information is displayed in a spatially limited area (Bier et al. 1993). For anatomical representations, this can allow different structures to be explored individually and in context (Wu et al. 2013).

In this work, a volumetric magic lens for AR applications is presented. It is designed for use in anatomical education and is intended to assist students in exploring the human body. An intuitive interaction is of particular interest to support the best possible way (Heinrich et al. 2020). Naturally, people interact with nearby objects with their hands, for example, by grasping, picking up, or moving them. This is also possible for virtual objects, with the option of both directly touching and more distant interaction (Valentini 2018). If the object must not be touched for manipulation, more people can use the visualisation at the same time from different viewpoints. However, if touching the object is not the starting point of the manipulation, other activation methods must be found. Therefore, this paper focuses on the investigation of different activation modes that define how mid-air hand gestures trigger object manipulations.

2. Related Work

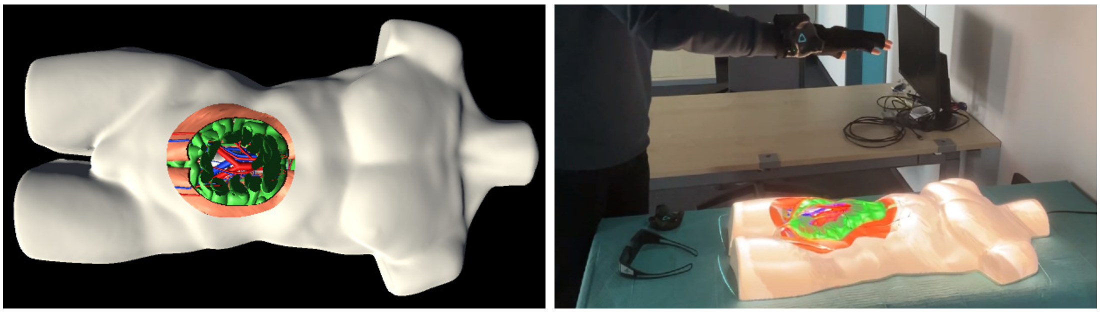
The magic lens concept was first introduced by Bier et al. (1993) for 2D visualisations. A magic lens applies a filter to the context display in a viewing region placed by the user. Filters may enhance or modify the view, displaying hidden or complementary information. Magic lenses are used in a wide range of applications, for example in art or scientific visualisations (Tominski et al. 2014). Several works discussed implementation options for magic lenses in 3D visualisations. Generally, a distinction can be made between flat and volumetric lenses (Viega et al. 1996). Flat lenses are 2D shapes in a 3D environment. These can be placed as freely movable planes (Viega et al. 1996), as well as be bound to the surface (Rocha et al. 2019). Virtual windows, in which hidden structures are shown through an opening in virtual objects (Bichlmeier et al. 2007), can be regarded as a special form of flat magic lenses (Monclús et al. 2009). Volumetric lenses are 3D objects that can be placed in the visualisation, where the filter is applied to a volume (Viega et al. 1996). They are often used to magnify or highlight structures that are otherwise hidden or hard to see (Traore et al. 2018). Pindat et al. (2013) combined a magnifying volumetric lens placed inside a multi layer object with a virtual window to highlight the position of the lens. Magic lenses are also used in a 3D AR environment which allows a more intuitive spatial placement of the viewing region compared to screen-based applications. In mobile AR, the display device itself might be used as a magic lens (Heinrich et al. 2017; Cao et al. 2007). Luo et al. (2021) combined 3D-objects displayed with an optical see-through head-mounted display, and an image slice view shown with a tablet used as a magic lens. Other handheld tools are also often used directly as magic lenses in AR, for example by projecting onto them (Spindler et al. 2010; Looser et al. 2007). Volumetric lenses are less common in AR. Monclús et al. (2009) presented a virtual magic lantern. There, a magnifying cone-shaped lens is placed with a handheld pointing tool. Mendez et al. (2006) investigated composable approaches with multiple volumetric lenses positioned with handheld markers. Heinrich et al. (2020) evaluated different interaction approaches to translate and scale a bounding box similar to a magic lens for direct volume rendering in spatial augmented reality. Foot and Hand gestures, as well as tracked handheld devices, were compared. Hand gesture interaction performed best in terms of task completion time and subjective workload. Recognition and integration of different gestures is also being explored specifically for AR anatomy training (Karambakhsh et al. 2019). Iqbal et al. (2021) compared a VR and an optical see-through AR setup as an alternative for cadaver dissection. They indicated difficulties in the gesture interaction with the AR glasses. Here only one gesture could be used. Blum et al. (2012) presented a virtual mirror for displaying medical volume data, in which slice images were controlled by gestures. However, most studies in the field focus on the anatomy education and visualisation, not on the gesture interaction. There are many works on touch-less interaction in the medical field in general due to sterility requirements (Besançon et al. 2021). However, oftentimes no interactions with 3D models are considered (Cronin and Doherty 2019). There, more degrees of freedom are required (Gallo 2013). In AR applications, hand recognition is often used to provide direct interaction with virtual models via pinch, grasp or grab gestures (Valentini 2018; Vuibert et al. 2015). For indirect 3D object manipulation using mid-air gestures, two hands might be used (Cui et al. 2016; Chen et al. 2017). In addition to hand recognition, body gestures are also utilised for touch-less interaction (Kirmizibayrak et al. 2011).

The activation of an interaction is particularly important. There are different approaches to activate the gesture recognition itself to avoid accidental detection, for example the use of gaze targets and trigger gestures by hand or foot (Hopmann et al. 2011; Hatscher and Hansen 2018). However, this has to be differentiated from the activation of the action associated with the gesture. Often, gestures are held for the duration of the interaction, and ending the gesture causes the action to end (Mentis et al. 2015). Another variant is the execution of an activation gesture, after which the action runs until an end gesture is detected (Heinrich et al. 2020). Other well known input methods like voice control, gaze interaction or foot gestures might be used instead or in addition. Mentis et al. (2015) evaluated voice commands and hand gestures for single actions, e.g. clicking a button, and manipulation modes, e.g. moving an object. Manipulation modes were controlled via voice control using discrete commands or via gestures using continuous manipulation. They recommend the redundant use of voice and hand interaction, but no evaluation of discrete commands versus continuous manipulation was provided.

3. Materials and Methods

3.1. The Magic Lens

A volumetric magic lens was used to create a spherical cut-out area inside the model, showing structures previously occluded (see Figure 1). A virtual window following the lens was added to ensure visibility of the cut area. It also allowed the user to look at the clipped scene from different angles and gives more flexibility by expanding the sphere with a cylinder shape. A similar approach with a cone shape was used by Pindat et al. (2013) to highlight the lens position and show occluded structures.



(a) The magic lens cuts a spherical shape out of the model, allowing the user to explore hidden structures.

(b) The lens position and size on the phantom of a human torso is controlled by mid-air hand gestures

Figure 1. The magic lens in the **(a)** development environment and **(b)** in spatial AR.

A Unity surface shader was developed to control how and if the vertices of the assigned object were rendered. The spherical lens was defined by a center point and a radius. Both parameters could be changed by the user to move and scale the lens. By default, the lens had a radius of 10 cm and was positioned on the stomach. A cylindrical extension of the lens was defined to generate the virtual window. The main axis of the cylinder was aligned towards the head position. The radius was determined by the scaling of the lens. For each vertex of an object assigned to the shader, the distance to the cutout area, consisting of the sphere and the cylinder, was calculated. All vertices within the area were clipped to generate a transparent lens. For each anatomical model, an individual material was generated based on the lens shader. This allowed an individual setting of render parameters for each model as well as a clear definition of objects to be cut through the lens.

3.2. Activation Modes

Activation modes define how the gestures had to be performed to start or stop the corresponding action. Two variants were examined in this work, based on discrete commands and continuous manipulation as described by Mentis et al. (2015). However, unlike Mentis et al., gesture interaction was used for both.

***Hold* activation mode:** In this mode, the corresponding action is executed as long as the gesture is held. The user has to hold a gesture when performing an action and takes a neutral hand position to stop the action.

***Toggle* activation mode:** In this mode, an action is started once by a gesture and then runs until it is stopped again by the same or another gesture. The user does not have to hold a gesture during the action. To end the action, the same gesture as that which was used for starting or that which was used for a universal stop gesture could be performed. Performing another start gesture during an active action led to the termination of the current and immediate start of the new action.

Only one activation mode could be active at a time. A change of the activation mode in the user study was done by the supervisor. The user was not able to switch between modes to avoid accidental changes.

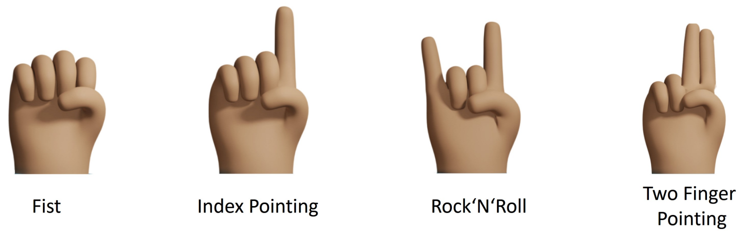


Figure 2. Illustrations of the hand gestures used to control the magic lens.

3.3. Gesture Control

Four mid-air hand gestures were used to interact with the magic lens (see Figure 2). Gesture control was implemented in a modular structure, from individual fingers to whole gestures. For each finger, the angle between fingertip, middle and base was calculated to determine if the finger was stretched (angle over 150 degrees), bent (angle under 100 degrees), or something in between. Gestures were then defined as state descriptions of the fingers. The thumb was excluded to avoid problems with gesture variations, for example, in the case of a fist with the thumb above or below the fingers. The sensitivity of the gestures was controlled by adjusting the bending and stretching angle threshold. In test runs prior to the study, one general setting worked well for all participants, so a calibration process was not needed. For a gesture to be recognised a second consecutive time, e.g. in toggle mode to stop an action, a neutral hand position had to be detected in between.

Fist Gesture for Translation: Translation of the lens was expected to be performed most often. Therefore, the gesture had to be easy to perform and hold. The fist gesture is not only easy but also intuitive, as it is a gesture resembling grabbing and moving an object. The lens cannot be grabbed directly as it is invisible and supposed to be inside the body. Therefore, the positional changes were implemented relative to the position of the hand when the gesture was activated. The user was, thus, able to control the lens position completely independently of its position. The lens followed the hand movements in three translational degrees of freedom, with a restriction of movement to close proximity to the phantom to avoid accidental misplacement. The fist gesture was activated when all fingers were recognised as bent.

Index Pointing for Scaling: The second important interaction for a spherical lens is scaling. To provide an intuitive metaphor, pointing with the index finger while moving the hand up or down, resembling usage of a slider, was used. The scaling was also performed relative to the position of the hand when the gesture was activated. This initial position for every activation was set to equal the current size of the lens. Scaling was restricted to a minimum radius of 2 cm and a maximum radius of 15 cm to prevent the lens from disappearing or getting too large, which rendered the lens unproductive. The Index Pointing was recognised when the index finger was stretched and the remaining fingers were bent.

Rock'N'Roll for Resetting: An interaction to reset the lens was needed in case of accidental misplacement or wrong scaling. It should not be activated by mistake and therefore required a more complex gesture. The Rock'N'Roll gesture is less intuitive but easy to remember as a pop culture reference. When activating the gesture, the lens was reset to a position on the stomach of the model in default size of 10 cm. The Rock'N'Roll gesture was recognised when the index finger and pinkie were stretched, with a bent middle and ring finger.

Two Finger Pointing for Stopping: The *Toggle* activation mode does not require the user to hold a gesture. This may result in the user forgetting which gesture they used for activation. A general stop gesture provides a reliable way to end all interactions immediately. Pointing with the index and middle finger is easy to perform and clearly distinguishable from the other gestures used. The Two Finger Pointing was recognised with a stretched index and middle finger and a bent ring finger and pinkie.

3.4. Setup

For gesture recognition, a data glove (Hi5 VR Glove, Noitom International Inc, USA) was used to enable the user to move freely around the model. An attached Vive tracker (HTC Corporation, Taiwan) provided the hand position. A bicycle helmet with a Vive tracker was used to track the head position, to enable perspectively correct rendering. The trackers and glove were connected to the game engine Unity (Unity Technologies, USA), which was used to implement the prototype. SteamVR (Valve Corporation, USA) and the Unity Plugin for the Hi5 VR Gloves (Noitom International Inc, USA) were used to integrate the tracking data. Two stereoscopic projectors (Barco G60-W7, Barco GmbH, Germany) with radio frequency 3D shutter glasses and a radio frequency emitter (VR 3D Glasses, Eyes3shut, France) were used with a torso phantom as the projection surface. With a photogrammetric measurement system (ProjectionTools, domeprojection.com GmbH, Germany), a 3D model of the phantom was created to generate and calibrate the projection surface mapping. Based on this and the head tracking data, spatially aligned and perspectively correct images were rendered onto the phantom with the associated Unity plugin (ProjectionTools, domeprojection.com, Germany). The organ models were retrieved from an anatomical database (Mitsuhashi et al. (2009): BodyParts3D, Japan) and were resized to fit the torso phantom.

3.5. Evaluation

A within-subject design user study was conducted to investigate the effects of the activation modes using the AR magic lens. The participants had to find a virtual sphere in the projected anatomy using the magic lens. The task was designed to promote frequent use and switching between gestures. Therefore, when the participant located the sphere, the lens had to be scaled to the size of the target. To reduce visual clutter, the organs and bones of the human body were shown in a semi-transparent blue while rendering the vessels and muscles transparent (see Figure 3). The target spheres were rendered green and were clipped by the magic lens like anatomical structures.

The activation mode was used as an independent variable with two levels: *Hold* and *Toggle*. Three dependent variables were measured. The *number of gesture changes* was used to evaluate if there was more switching between gestures in a mode. The *task completion time* was used to check whether the task could be completed faster in one activation mode than in the other. To determine if the task could be better accomplished with one mode, the correctness of task performance was ascertained. To measure the *accuracy*, two metrics were used to evaluate the overlap of the lens and target spheres. The Jaccard index was used to evaluate similarity. Both the scale and the position in space should be taken into account. A comparison of the volumes alone would not have given sufficient information about the overlap. To evaluate whether the target was found with the lens, while ignoring the correct scaling, the overlap coefficient was used additionally. The *subjective assessment* was collected using a Likert scale

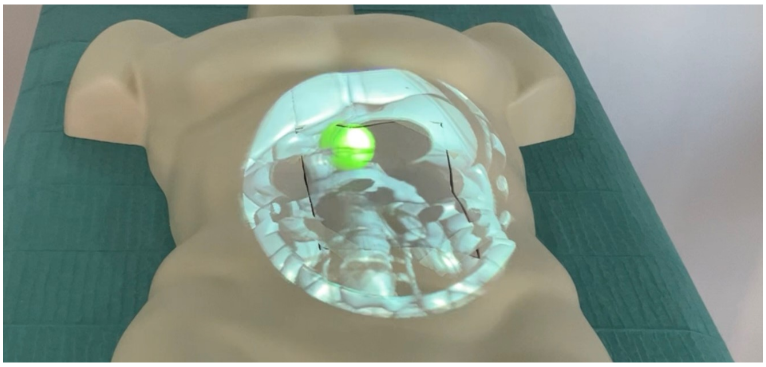


Figure 3. Modification of the magic lens to reduce visual clutter in the user study.

based on the NASA Task Load Index questions. Participants were asked to rate the mental and physical demand, the perceived time pressure, necessary effort, and stress during the task on 5-point Likert items ranging from 1 (*very low*) to 5 (*very high*). The *subjective assessment* was retrieved by adding up the scores of all Likert items, resulting in a Likert score ranging from 0 to 25.

Because no advanced prior medical or technical knowledge was needed for the tasks, the study was advertised in local and university online groups without recruitment requirements. Twenty-four participants (11 male, 11 female, 2 diverse), aged from 20 to 41 years (median: 24 years) took part in the study. Most had little experience with either AR or other mixed reality applications, or hand gestures and human anatomy. Each participant received 10€ as compensation.

First, written consent was collected for each participant. Then, the setup, interaction and study task were explained. Participants were then asked to put on the stereo glasses, the head tracking helmet and the gloves for the gesture recognition. The first activation mode was thereafter explained in a tutorial, allowing subjects to familiarise themselves with the interaction methods, gestures and the task. Additionally, a reference sheet for the gestures was available at all times. The participants were then instructed to favour accuracy over speed and to take as much time as needed. Subsequently, four tasks were performed, using the first activation mode. Afterwards, a tutorial of the second activation mode was presented. Again, participants were allowed to test out the interaction methods in a sample task. Four task repetitions with the second activation mode then followed. For each activation mode, the Likert scale was completed and qualitative feedback was collected. After completing all tasks, the participants were asked to fill out a questionnaire, retrieving demographic information as well as user feedback about the magic lens, interaction and setup.

The starting activation mode was alternated between participants to avoid learning effects. Additionally, eight predefined target spheres with varying positions in the upper body and sizes varying between 4 cm to 10 cm were randomly assigned to the task for this purpose.

4. Results

After completion of the study, the Jaccard index, Overlap coefficient, and the number of gesture changes with and without the Two Finger Pointing stop gesture were determined for all trials. The *subjective assessment* score was retrieved by summing up

Likert items. Subsequently, mean values were determined for measurement repetitions of identical experimental conditions for each participant.

The acquired data were tested for normal distribution using Shapiro-Wilk tests. A normal distribution was not given for all groups. Since the requirements for a paired t-test were therefore not met, a paired Wilcoxon signed rank test was applied to evaluate the statistical significance of the results. The effect of the activation modes is presented in Figure 4.

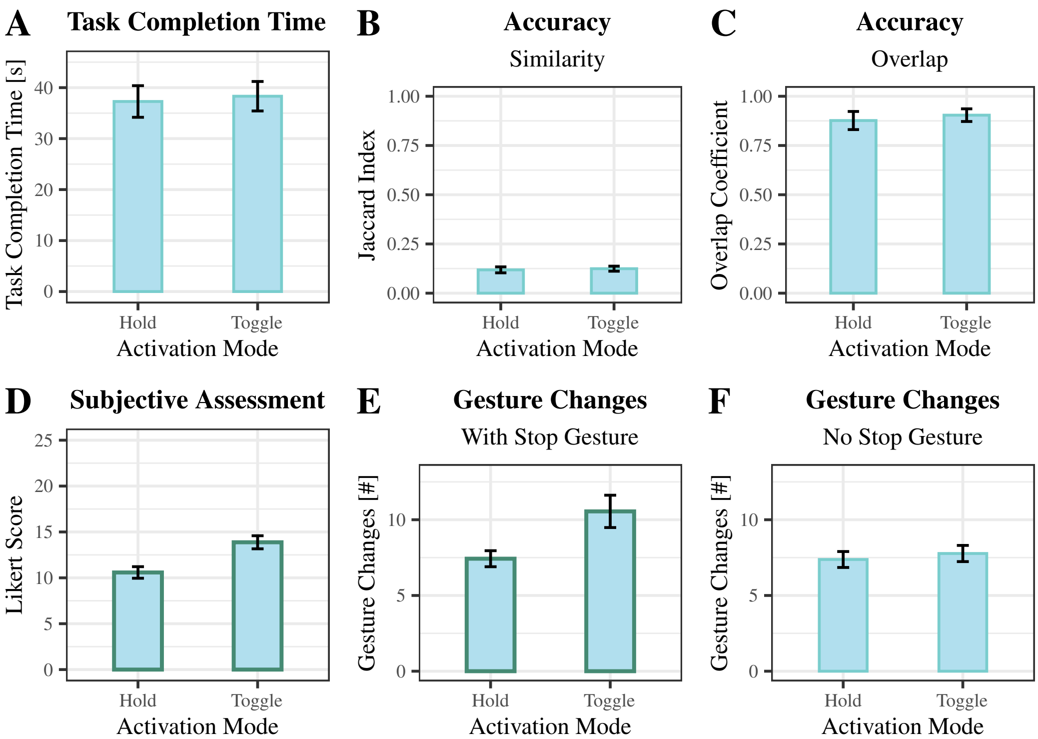


Figure 4. Effects of the activation mode on the **A** *task completion time*, **B** & **C** *accuracy* using the Jaccard index to evaluate the similarity of the target and lens volume, and the Overlap coefficient to assess to which extent the lens is within the target (both ranging from 0 to 1), **D** *subjective assessment* based on a Likert scale (ranging from 0 to 25), **E** & **F** *number of gesture changes* with and without the stop gesture. Bold, darker outlines indicate significant results in the Wilcoxon signed rank test.

Regarding the *number of gesture changes*, significantly more changes were recorded for the *Toggle*-Mode than with the *Hold*-Activation ($z = -2.72$, $p = 0.006$). However, when leaving out the Two Finger Pointing as a stop gesture, which was only needed in the *Toggle*-Mode, no significant differences were found for the interaction count ($z = -0.55$, $p = 0.583$). This shows that there was no increased switching between the different interactions, even if more gestures were used. It was noticed in the study that the manner of stopping (with the stop gesture, with repetition of the active gesture or with execution of another action gesture) differed from user to user. Only small differences were found for the *task completion time* ($z = -0.55$, $p = 0.584$) and the *accuracy* (Jaccard index: $z = -0.01$, $p = 0.989$; Overlap coefficient: $z = -0.54$, $p = 0.587$). The activation modes thus did not influence how well the task could be completed. Overall, the Jaccard index for the *accuracy* turned out to be very low. This suggests a task that was challenging for many participants. Another explanation can be found in the task design. The sphere was supposed to be scaled to the same size as the target. However, a lens of exactly the same size and position made it impossible to see the target and therefore confirm the positioning. Users thus often scaled the lens slightly smaller than the target, or misplaced it somewhat. The overlap coefficient shows that the lens, as the smaller volume, was always completely or mostly inside the target. Differences in volume are strongly penalised in the Jaccard index and therefore resulted in low scores.. Finding the target was possible with the lens, but scaling to fit the target was difficult. This may also be due to problems with depth perception that were frequently reported in the qualitative feedback. The difficulty can also have had an influence on the *task completion time*. Since the task itself took a long time, differences caused by the activation modes may be masked. The overall experience was rated positively. The gestures were overall described as easy to perform with low to medium difficulty to recall. Participants reported no ergonomic problems with the setup, however some technical issues, e.g. jittering or gestures that were not immediately recognised, were reported. *Hold* performed significantly better in the *subjective assessment* ($z = -3.4$, $p < 0.001$). It was described as less mentally and physically demanding with a higher success rate. Most participants felt that the *Toggle* mode was harder to understand and counter-intuitive. It can be concluded that the different activation modes do not influence the final result, but do have an effect when performing the task.

5. Discussion

The *Hold* mode is favoured in the *subjective assessment* as well as additional qualitative feedback. While both activation modes performed similar in terms of *task completion time* and *accuracy*, more gesture changes were needed due to an additional stop gesture in the *Toggle* mode. However, our study was limited to a small selection of gestures that were easy to perform and hold. Future study designs should also include more complex gestures that my be harder to hold. Furthermore, the implementation is currently limited to static gestures. Dynamic gestures, like swiping, may not only yield different results, but require new activation modes. This should be considered in future studies. Additionally, no handedness was considered in our work. This was not criticised by any participant, but in the future, it should be considered whether the use of the left or both hands has an effect compared to the right hand only. This work considers hand gesture control using gloves to enable a large recognition area. Future work might explore solutions requiring no gloves to increase the system ergonomics .

The results for *accuracy* indicates problems in perceiving the lens. The Jaccard index and, thus, the similarity between the spheres was low. At the same time, the overlap coefficient was very high, the lens was almost always completely in the target. This is partly caused by difficulties with depth perception, even in stereoscopic projection systems (Heinrich et al. 2022) and by the visualisation itself. The magic lens is transparent, which makes it inherently elusive and hard to grasp for the user. This also caused difficulties for some users trying to differentiate scaling of the lens from translation, as both are able to increase or decrease the depth of the cutout into the body. Another difficulty caused by the transparency was the inability of many participants to differentiate between the lens itself and the cylinder between user and lens, which was also made transparent to prevent the lens from getting lost in the body. Further ways of enhancing the visibility of the lens may be investigated but should not unnecessarily restrict the visibility of the model itself. Here, interactive bounding boxes (Heinrich et al. 2020) or outlines for the virtual window (Pindat et al. 2013) might be helpful. In addition, a different task should be used in future studies. Our task led to an underscaling and thus low *accuracy* of the lens in almost all subjects. Other task ideas could, for example, use more abstract size specifications in the form of diameter measurements if scaling is to be enforced.

The task led to long periods of interaction without gesture changes, for example only translation when searching the target. In addition, participants took a long time to complete the task, presumably due to its difficulty. The complexity and required time, especially whilst staying in one activation mode, might have marginalised the quantitative effects of the evaluated activation modes. Investigating other tasks could be interesting here. Independent factors on task completion, like in which order participants search the body in a search task like the one used in our study, should be avoided. Task designs with better quantifiable accuracy of the results are also preferable. In particular a differentiation between broad and fine adjustments in the lens manipulation should be investigated. The user might have other interaction requirements for low precision tasks, than for more accurate ones. Experiments with multiple parallel tasks might be better suited to estimating and comparing the cognitive load required for activation modes, as well as simulating a situation closer to real use scenarios. Our task design did not require participants with medical knowledge, as we evaluated two general activation modes. Choosing medical students as participants would have additionally allowed to evaluate anatomy-specific interaction and visualisation aspects. This might be explored in future work. More realistic models could be more appropriate for this tasks, as well as an evaluation of the lens for anatomy specific tasks, eg. finding and exploring specific structures of the body.

When measuring both time and accuracy, their interdependence has to be considered. Participants were instructed to favour accuracy over time, therefore taking longer to finish.As our study was a within-subject-design, paired measurements were investigated, making these dependencies less relevant. Therefore, we do not believe that this had a negative impact on our results.

6. Conclusion

We implemented a Magic Lens controlled by mid-air hand gestures for the interactive exploration of human anatomy. Our evaluation of two different gesture activation modes revealed the *Hold* activation mode to be preferable. Qualitative feedback revealed a significant preference for the *Hold*-mode by the users. No significant differences were found for the user performance in the task. However, more interactions are needed in the *Toggle* mode due to an additionally required stop gesture. Whilst an anatomical model was used for the study, the tasks itself did not require medical knowledge, but instead focused on comparing the activation modes. This makes the results of our study generalisable to other application areas. Extensions such as searching, highlighting and other customisations are also interesting for use in medicine. Further studies might explore how the lens can be used best for anatomy learning. More research is needed into which gestures are best suited for which interaction. Aligning the gestures closer to the metaphor of a lens is an interesting approach to gesture interaction that may be explored in future work. Other methods of interaction, like voice control, may be promising.

Acknowledgements

This work was partially funded by the the German Research Foundation under Grant HA 7819/1-2 within the STIMULATE research campus; the German Federal Ministry for Economic Affairs and Climate Action under Grant 16KN093921. The authors report there are no competing interests to declare.

References

Besançon L, Ynnerman A, Keefe DF, Yu L, Isenberg T. 2021. The state of the art of spatial interfaces for 3d visualization. Computer Graphics Forum. 40(1):293–326.

Bichlmeier C, Wimmer F, Heining SM, Navab N. 2007. Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. IEEE. p. 129–138.

Bier EA, Stone MC, Pier K, Buxton W, DeRose TD. 1993. Toolglass and magic lenses. In: Whitton MC, editor. Computer graphics proceedings; New York, New York, USA. ACM Press. p. 73–80. Available from: https://dl.acm.org/doi/pdf/10.1145/166117.166126.

Bjork S, Holmquist LE, Redstrom J. 1999. A framework for focus+context visualization. In: Wills G, Keim D, editors. 1999 IEEE symposium on information visualization (InfoVis'99). IEEE Comput. Soc. p. 53–56.

Blum T, Kleeberger V, Bichlmeier C, Navab N. 2012. mirracle: An augmented reality magic mirror system for anatomy education. In: Author IVRCC, editor. 2012 IEEE Virtual Reality. IEEE. p. 115–116.

Buchner J, Buntins K, Kerres M. 2022. The impact of augmented reality on cognitive load and performance: A systematic review. Journal of Computer Assisted Learning. 38(1):285–303.

Cao X, Forlines C, Balakrishnan R. 2007. Multi-user interaction using handheld projectors. In: Shen C, Jacob R, Balakrishnan R, editors. Proceedings of the 20th annual ACM symposium on User interface software and technology - UIST '07; New York, New York, USA. ACM Press. p. 43.

Chen LC, Cheng YM, Chu PY, Sandnes FE. 2017. Identifying the usability factors of mid-air hand gestures for 3d virtual model manipulation. In: Antona M, Stephanidis C, editors. Universal Access in Human–Computer Interaction. Designing Novel Interactions; Cham. Springer International Publishing. p. 393–402.

Cronin S, Doherty G. 2019. Touchless computer interfaces in hospitals: A review. Health informatics journal. 25(4):1325–1342.

Cui J, Kuijper A, Fellner DW, Sourin A. 2016. Understanding people's mental models of mid-air interaction for virtual assembly and shape modeling. In: Proceedings of the 29th International Conference on Computer Animation and Social Agents; New York, NY, USA. Association for Computing Machinery. p. 139–146. CASA '16.

Estai M, Bunt S. 2016. Best teaching practices in anatomy education: A critical review. Annals of anatomy = Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft. 208:151–157.

Gallo L. 2013. A study on the degrees of freedom in touchless interaction. In: SIGGRAPH Asia 2013 Technical Briefs; New York, NY, USA. Association for Computing Machinery. p. 1–4. SA '13.

Ghosh SK. 2017. Cadaveric dissection as an educational tool for anatomical sciences in the 21st century. Anatomical sciences education. 10(3):286–299.

Hatscher B, Hansen C. 2018. Hand, foot or voice. In: D'Mello SK, Georgiou P, Scherer S, Provost EM, Soleymani M, Worsley M, editors. Proceedings of the 20th ACM International Conference on Multimodal Interaction; New York, NY. ACM. p. 145–153.

Heinrich C, Langlotz T, O'Keefe R. 2017. [poster] volume lens: Exploring medical volume datasets using mobile devices. In: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct). IEEE. p. 35–38.

Heinrich F, Bornemann K, Lawonn K, Hansen C. 2020. Interacting with medical volume data in projective augmented reality. In: Martel A, editor. Medical image computing and computer assisted intervention – miccai 2020. (LNCS Sublibrary: SL6 - Image Processing, Computer Vision, Pattern Recognition, and Graphics; vol. 12263). Cham, Switzerland: Springer; p. 429–439.

Heinrich F, Schwenderling L, Joeres F, Hansen C. 2022. 2d versus 3d: A comparison of needle navigation concepts between augmented reality display devices. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE. p. 260–269.

Hopmann M, Salamin P, Chauvin N, Vexo F, Thalmann D. 2011. Natural activation for gesture recognition systems. In: Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11; New York, New York, USA. ACM Press. p. 173–183.

Iqbal MZ, Mangina E, Campbell AG. 2021. Exploring the real-time touchless hand interaction and intelligent agents in augmented reality learning applications. In: Economou D, editor. Proceedings of 2021 7th International Conference of the Immersive Learning Research Network (iLRN); Piscataway, NJ. IEEE. p. 1–8.

Karambakhsh A, Kamel A, Sheng B, Li P, Yang P, Feng DD. 2019. Deep gesture interaction for augmented anatomy learning. International Journal of Information Management. 45:328–336. Available from: https://www.sciencedirect.com/science/article/pii/S0268401217308678.

Kirmizibayrak C, Radeva N, Wakid M, Philbeck J, Sibert J, Hahn J. 2011. Evaluation of gesture based interfaces for medical volume visualization tasks. In: Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry - VRCAI '11; New York, New York, USA. ACM Press. p. 69–74.

Looser J, Grasset R, Billinghurst M. 2007. A 3d flexible and tangible magic lens in augmented reality. In: 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007; Piscataway, NJ. IEEE. p. 51–54.

Luo W, Goebel E, Reipschlager P, Ellenberg MO, Dachselt R. 2021. Exploring and slicing volumetric medical data in augmented reality using a spatially-aware mobile device. In: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). IEEE. p. 334–339.

Mendez E, Kalkofen D, Schmalstieg D. 2006. Interactive context-driven visualization tools for augmented reality. In: IEEE/ACM International Symposium on Mixed and Augmented Reality, 2006; Piscataway, NJ. IEEE Service Center. p. 209–218.

Mentis HM, O'Hara K, Gonzalez G, Sellen A, Corish R, Criminisi A, Trivedi R, Theodore P. 2015. Voice or gesture in the operating room. In: Begole B, editor. CHI 2015 : extended abstracts publication of the 33rd annual CHI Conference on Human Factors in Computing Systems : April 18-23, 2015, Seoul, Republic of Korea. ACM. p. 773–780.

Mitsuhashi N, Fujieda K, Tamura T, Kawamoto S, Takagi T, Okubo K. 2009. BodyParts3D: 3D structure database for anatomical concepts. Nucleic Acids Research. 37:D782–D785.

Monclús E, Díaz J, Navazo I, Vázquez PP. 2009. The virtual magic lantern. In: Kitamura Y, Takemura H, Kiyokawa K, Lok B, Thalmann D, Spencer SN, editors. Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology - VRST '09; New York, New York, USA. ACM Press. p. 119.

Pindat C, Pietriga E, Chapuis O, Puech C. 2013. Drilling into complex 3d models with gimlenses. In: Thalmann NM, Wu E, Tachi S, Thalmann D, Nedel L, Xu W, editors. Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology - VRST '13; New York, New York, USA. ACM Press. p. 223.

Rocha A, Silva JD, Alim UR, Carpendale S, Sousa MC. 2019. Decal-lenses: Interactive lenses on surfaces for multivariate visualization. IEEE transactions on visualization and computer graphics. 25(8):2568–2582.

Spindler M, Tominski C, Schumann H, Dachselt R. 2010. Tangible views for information visualization. In: ACM International Conference on Interactive Tabletops and Surfaces - ITS '10; New York, New York, USA. ACM Press. p. 157–166.

Tominski C, Gladisch S, Kister U, Dachselt R, Schumann H. 2014. A survey on interactive lenses in visualization. In: EuroVis (STARs). The Eurographics Association. p. 43–62.

Traore M, Hurter C, Telea A. 2018. Interactive obstruction-free lensing for volumetric data visualization. IEEE transactions on visualization and computer graphics.

Valentini PP. 2018. Natural interface for interactive virtual assembly in augmented reality using leap motion controller. International Journal on Interactive Design and Manufacturing (IJIDeM). 12(4):1157–1165.

Viega J, Conway MJ, Williams G, Pausch R. 1996. 3d magic lenses. In: Kurlander D, Brown M, Rao R, editors. UIST '96; New York. ACM. p. 51–58.

Vuibert V, Stuerzlinger W, Cooperstock JR. 2015. Evaluation of docking task performance using mid-air interaction techniques. In: Banic A, editor. SUI '15 : proceedings of the 3rd ACM Symposium on Spatial User Interaction : August 8-9, 2015, Los Angeles, CA, USA; [Place of publication not identified]. ACM. p. 44–52.

Wu ST, Vidalon JEY, Loos WS, Coan AC. 2013. Query tools for interactive exploration of 3d neuroimages: Cropping, probe and lens. In: 2013 XXVI Conference on Graphics, Patterns and Images. IEEE. p. 250–257.

Yammine K, Violato C. 2015. A meta-analysis of the educational effectiveness of three-dimensional visualization technologies in teaching anatomy. Anatomical sciences education. 8(6):525–538.