3D spacecraft configuration using immersive AR technology

Artur Baranowski*, Sebastian Utzig*, Philipp Fischer*, Andreas Gerndt*, Jens Herder†

* German Aerospace Center
Lilienthalplatz 7
38108 Braunschweig

† Hochschule Düsseldorf
Münsterstraße 156
40476 Düsseldorf

Abstract: In this paper we propose an integrated immersive augmented reality solution for a software tool supporting spacecraft design and verification. The spacecraft design process relies on expertise in many domains, such as thermal and structural engineering. The various subsystems of a spacecraft are highly interdependent and have differing requirements and constraints. In this context, interactive visualizations play an important role in making expert knowledge accessible. Recent immersive display technologies offer new ways of presenting and interacting with computer-generated content. Possibilities and challenges for spacecraft configuration employing these technologies are explored and discussed. A user interface design for an application using the Microsoft HoloLens is proposed. To this end, techniques for selecting a spacecraft component and manipulating its position and orientation in 3D space are developed and evaluated. Thus, advantages and limitations of this approach to spacecraft configuration are revealed and discussed.

Keywords: spacecraft configuration, human computer interaction, augmented reality, 3D user interfaces, HoloLens

1 Introduction

The development of a spacecraft is a complex interdisciplinary endeavor, requiring expertise in many domains, such as communications, attitude and orbit control as well as structural and thermal engineering. To meet the requirements and constraints of each subsystem, interdependencies have to be understood and balanced. Thus, reliable means of communication are an integral part of the spacecraft design process. The German Aerospace Center (DLR) employs a Concurrent Engineering (CE) approach to support collaboration between domain experts [FDM+17]. Within the Concurrent Engineering Facility (CEF) in Bremen, co-located engineers can discuss issues in a face-to-face manner facilitating the exchange of information and data. Using the integration software Virtual Satellite (VirSat), data consistency is ensured during a design session. In VirSat, properties of a spacecraft and its subsystems relevant to early design stages can be modelled, thus providing engineers with a common shared data model. However, domain experts are still facing the challenge of expressing their specialized knowledge to engineers of other subject areas [FWG12]. Especially configuration issues concerning the position and orientation of spacecraft components prove to be difficult to communicate. To address this issue, VirSat was extended by additional configuration parameters used to store visualization data. Each system component can be associated with a geometric primitive, color and transparency level. Also included are position and orientation parameters of a component in 3D space with respect to its parent component. Configuration parameters are manipulated via UI widgets, such as text input fields or drop down menus (see figure 1). Moreover, a view to display the 3D visualizations has been included (figure 2).

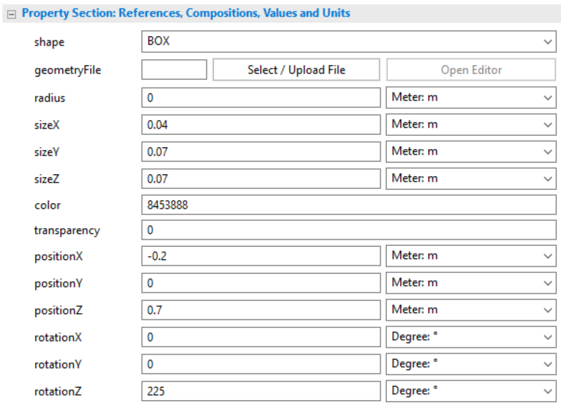

Figure 1: 2D user interface for configuration parameter input.

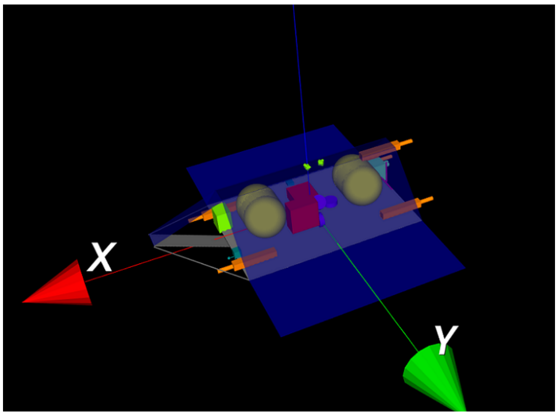

Figure 2: 3D preview of a spacecraft model based on VTK.

Advancements in Virtual Reality (VR) and Augmented Reality (AR) offer new ways of presenting computer-generated content. Head-Mounted Displays (HMDs), such as the Microsoft HoloLens, allow for a perspective view that is coupled to the viewpoint of the user. Stereoscopy is enabled by projecting 3D content via small near-to-eye displays. Ware et al. provide evidence that these properties aid the understanding of complex virtual models [WAB93]. Furthermore, the HoloLens provides advanced input modalities such as voice or gesture control, which allow for direct and intuitive interaction within virtual environments. Thus, benefits of employing such a device can be twofold: both the understanding of a satellite construction and the ease of configuring its parameters can be improved. Moreover, the HoloLens is a fully mobile device enabling engineers to move freely within their workspace. With the see-through visor of the HoloLens natural communication between peers remains intact. Additionally, engineers are able to use the VirSat 2D interface in combination with a 3D interface provided by HoloLens.

Our aim is to reveal advantages and limitations of spacecraft configuration using immersive AR technology. This is achieved by designing a tailored user interface for the HoloLens and evaluating it against the VirSat interface. Hereby, the core objective lies in providing efficient and accurate means for manipulating six degrees of freedom (6DoF). Specifically, this means configuring a component's position and rotation with respect to three dimensions in space. We begin by giving an overview of related work in section two. In section three the technical framework of the HoloLens is briefly described. Section four explains the developed

HoloLens user interface we evaluated. Results of this evaluation are presented in section six, following a discussion. Finally, conclusions are drawn and an outlook for further research opportunities is given.

2 Related Work

Spacecraft configuration using immersive AR technology stands to benefit from plenty of research conducted in the realms of 3D manipulation and 6DoF input techniques. On a general level, important design guidelines can be derived from academic literature. According to Bowman et al. the number of control dimensions (degrees of freedom) and their integration (simultaneous manipulation) are key factors regarding 3D manipulation [BKLP04]. Zhai et al. determined that integrated control of multiple input dimensions usually results in superior user performance [ZMR97]. However, regarding mid-air interaction techniques commonly used for immersive devices such as the HoloLens, separation of transformations yielded better results [MRFJ16]. By separating the rotational and translational components, unwanted manipulations can be avoided.

This is especially important, because it is cognitively more demanding to dissect object rotations than object translations. In user studies Frees et al. observed participants having difficulties in performing purposeful rotations and finding an accurate axis of rotation [FKK07]. Likewise, Parsons recognized a general inability to reason about an object's rotation in terms of angles and axes [Par95]. However, Frees et al. observed that once participants found an accurate rotational axis they were able to solve the posed tasks with ease. Considering these observations, integrated control of all three rotational input dimensions seems especially important for enabling users to quickly determine the desired rotational axis.

Furthermore, limitations in input precision, either due to human accuracy or low input device resolution, pose a problem when performing fine-grained manipulations. Moreover, limited tracking areas constrain interaction space, making it difficult to cover longer distances without losing tracking. A common approach addressing these issues is motion scaling, as proposed by Frees et al. [FKK07] with the PRISM (Precise and Rapid Interaction through Scaled Manipulation) interaction technique. Derived from Fitts' Law [Fit92], a need for precision is implied by slow hand movements, whereas fast hand movements indicate "rough" approximate manipulations. PRISM exploits these observations to scale object movement depending on hand velocity, specifying the magnitude of scaling with seamless transition between downscaled and unscaled object movements. Constraint-based manipulations are another way to improve input precision. By using handles co-located to the manipulated object, a particular degree of freedom of the object's transform is individually manipulated. However, the increase in placement precision comes at the expense of speed, especially for complex tasks [MRFJ16]. The 3-Point++ tool is another approach enabling precise manipulation using handles [ND13]. This tool explicitly defines an object's position and rotation by three handles and their barycenter.

Based upon a classification by Bowman et al. [BKLP04], four basic tasks can be regarded as fundamental for 3D manipulation:

- **Selection:** Identification and acquisition of a component from a set of components.
- **Positioning:** Changing a targets position in 3D space.
- **Rotation:** Changing a targets orientation in 3D space.
- **Scaling:** Changing an objects extensions in 3D space.

This decomposition simplifies 3D manipulations to their most essential constituents and serves as a blueprint for an initial interface design for spacecraft configuration.

3 Technical Background

The adequacy of interaction techniques is highly dependent on the underlying technological framework. Hence, certain capabilities and limitations of the HoloLens are crucial for the user interface design. Running four "environment-sensing cameras" and an inertial measurement unit, it can track its position and orientation in space [Col18]. In doing so, a "gaze-based ray-casting" (Gaze) metaphor can be employed by extending a ray in view direction. In addition, a depth camera enables gesture recognition and tracking of hand position in space. However, hands can only be tracked within the approximately 120×120 degree wide field of view of the camera. Furthermore, hand tracking is limited to positional parameters. Rotational hand parameters are not yet captured. A possible way to solve this issue is to extend tracking capabilities by introducing external input devices, such as the Leap Motion [Mot18]. While also having a limited tracking area, palm rotation could be mapped to an object's orientation. However, we believe that one of the HoloLenses most valuable assets is its mobility gained by inside-out tracking techniques. By using external devices which are either tethered, reliant on outside-in tracking techniques or remote workstations mobility is compromised. Therefore, the input system provided by the HoloLens will be used exclusively.

3.1 Gesture Recognition

Two distinct hand states can be recognized by the HoloLens: the "ready" and the "pressed" state (see figure 3). All core gestures are built upon these hand states [Cen18]. Rapidly transitioning from ready into pressed state constitutes a so-called "Air-Tap". In conjunction with Gaze for targeting virtual objects or UI elements, this motion sequence resembles the point-and-click behavior commonly known from generic desktop interfaces. A "Double-Tap" is composed of two successive "Air-Taps". Both constitute discrete actions, whereas a "Tap-and-Hold" gesture can be used to perform continuous manipulations, such as moving an UI window. A "Tap-and-Hold" gesture is active for as long as the hand remains in pressed state.

Figure 3: “Ready” (left) and “pressed” (right) hand pose.

3.2 Virtual Satellite Architecture

The HoloLens application is integrated into the architecture of Virtual Satellite and continuously receives updates on visualization parameter changes from a VirSat-Server-Instance. On the other hand, client-side manipulations on the HoloLens are validated by the server and applied to the shared data model. This way, client and server instances always have a synchronized spacecraft model.

4 Interaction

The primary requirement of the proposed user interface is to provide suitable techniques for performing the essential tasks of 3D interaction. Object selection is based upon a multimodal input technique. Using Gaze, an object is targeted by a cursor centered in the user’s field of view. This way, gestural input can be associated to a particular component. The targeted object is then selected by performing an “Air-Tap”. To give the user visual feedback about the selection, we use a magenta bounding box to enclose the selected object (see figure 4 and 5).

A selected object can be transformed. The depth camera can track both hands 3D position in space, equating integrated control along six input dimensions. As determined by Mendes et al. [MRFJ16] however, separating manipulation of an object’s translational and rotational components resulted in improved user performance. Therefore, transformations are separated into two exclusive modes, indicated by either cubic handles (translation mode) or spherical handles (rotation mode) on the corners of the bounding box. Translation mode enables the user to manipulate a component’s position in 3D space. During a “Tap-and-Hold” gesture, the positional offsets between two consecutive hand positions are continuously measured and added to a target components position. This equates to a position control technique imitating real-world interactions with a physical object.

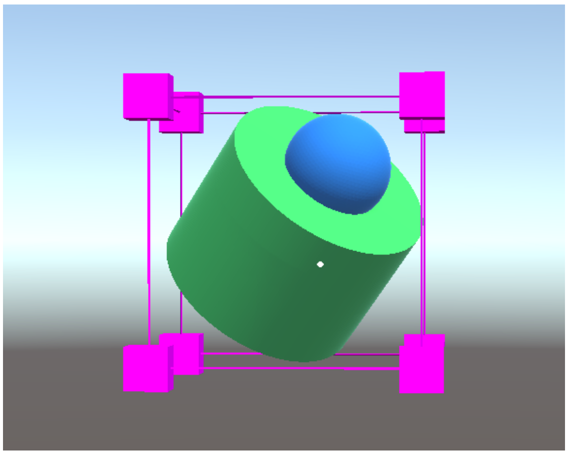

Figure 4: Cubic handles indicating object is in translational mode.

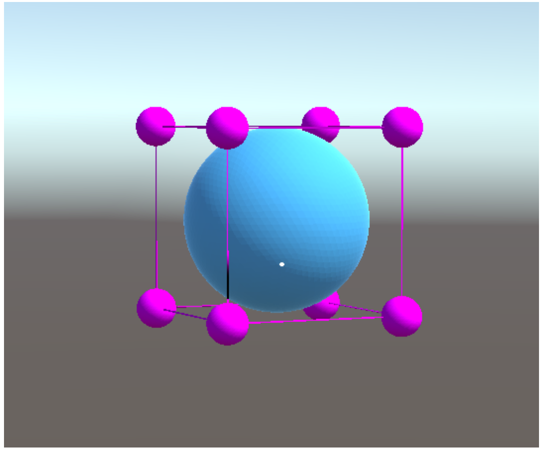

Figure 5: Spherical handles indicating object is in rotational mode.

Under the premise that rotational parameters are not captured by the HoloLens, finding a suitable metaphor for manipulating an object's rotation poses a challenge. Generally, constraint- and handle-based approaches enable reliable and precise configuration of an object's rotation [MRFJ16] [ND13]. However, with these techniques' manipulation speed is sacrificed. Moreover, these approaches are more indirect and very different from the direct translation mode manipulations. As explained in section 2, a general inability to reason about an object's rotation in terms of angles and axes impedes making purposeful rotations when manipulating an object [Par95]. Based on these considerations, we want to provide integrated control over an object's rotation, enabling users to intuitively find a rotational axis by observing an object's rotational behaviour. This is achieved by mapping the displacement of a user's hand to changes in the component's orientation in rotation mode. Similar to translation mode, offsets between two consecutive hand positions are added to a component's rotation. Unlike in translation mode however, these offsets are transformed into the user's view coordinate system. This way, object rotation is coupled to the user's perspective, ensuring similar rotation behavior from any angle of view.

4.1 Motion scaling

The alignment of specific spacecraft components can be sensitive to input precision. For example, successful attitude control highly depends on a precise arrangement of satellites reaction wheels. Moreover, the depth cameras limited field of view constrains interaction space, making it difficult to cover longer distances without losing tracking. Thus, techniques for improving input precision and extending interaction space are necessary. Based on these considerations, a motion scaling technique similar to the PRISM interaction technique proposed by Frees et al. [FKK07] is implemented. The implicit switch between scaled and unscaled mode puts no additional mental overhead on the user. Moreover, the technique can easily be combined with more sophisticated manipulation metaphors based on constrained manipulation or handles in future development iterations.

The ratio between an input quantity (hand motion) and an output quantity (target object motion) is manipulated by multiplying it with a scaling factor (see figure 6). Slow hand motion, indicating a need for precision, is downscaled resulting in even slower object motion. Likewise, fast hand motion translates to upscaled object motion. A scaling factor is determined according to three constants. Hand velocities below a minimal velocity are suppressed, thus filtering involuntary minuscule movements. Hand velocities below a downscaling constant are scaled by a factor that diminishes quadratically in proportion to hand velocity. Likewise, hand velocities above an upscaling constant are scaled by a factor that increases quadratically in proportion to hand velocity. In between hand motion is not scaled and mapped one-to-one to object motion.

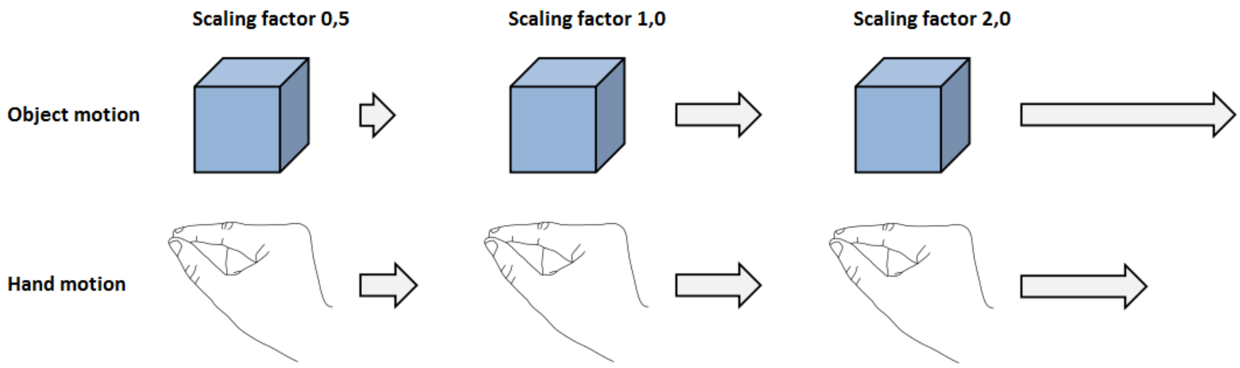

Figure 6: Slow hand motion halving object motion by a factor of 0.5 (left). Hand motion mapped to object motion one-to-one (middle). Fast hand motion doubling object motion by a factor of 2 (right).

5 Evaluation

A user study was conducted to compare the VirSat 2D interface with the new HoloLens 3D interface in context of satellite configuration. The following research hypotheses were established:

1. Difference in task completion time between HoloLens and VirSat is significant.
2. Difference in positional error between HoloLens and VirSat interface is significant.
3. Difference in angular error between HoloLens and VirSat interface is significant.

Twelve participants took part in a within-group study. All participants were technically affine and had prior experience with generic desktop interfaces based on common UI widgets like sliders, drop down menus, buttons or text input fields. Ten participants had prior experience using 3D graphics or CAD tools such as Unity3D, Blender or CATIA. Four had prior experience using the HoloLens.

Participants performed three 3D docking tasks with both the HoloLens and VirSat interface. These tasks required the participants to select, place, and rotate an object with respect to a target transformation (see figure 7). Before each task, participants were shown an instructional video explaining the task. The aim was to mimic work processes during a design study in the CEF in which peers have to explain the needs of their respective subsystems to each other. The participants were shown the initial and target configuration (see figure 7). The tasks required the participants to place a cylindric shape to a particular side of a cubic shape. To counter-balance fatigue and learning effects, task and interface order were randomized. Before each trial participants were given ten minutes to familiarize themselves with each interface by performing a training task.

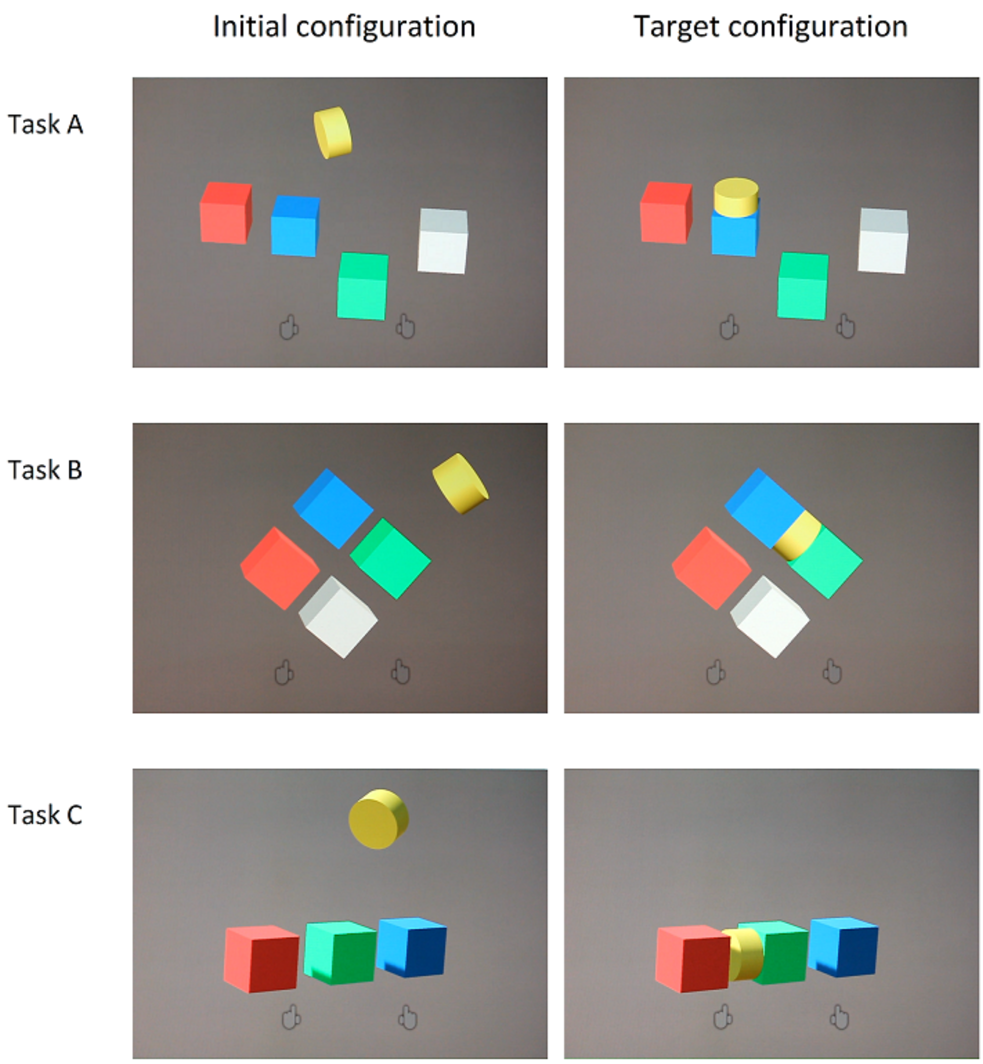

Figure 7: Configuration tasks performed during trials: initial configuration (left) and target configuration (right).

6 Results

Completion time per task averaged 139.6 seconds for the HoloLens interface and 187.2 seconds for the VirSat interface (see figure 9). This is equal to a mean difference of 47.6 seconds. Figure 8 shows the distribution of both data sets in a box-and-whisker diagram. The median value is 99.5 seconds for the HoloLens interface. Upper and lower quartile values are 192.0 and 75.5 seconds. For the VirSat interface median value is 161.0 seconds. Upper and lower quartile values are 239.0 and 121.25 seconds. The measurements are normally distributed, as a Kolmogorov-Smirnov test for normality shows. Furthermore, a paired t-test verifies statistical significance regarding the difference between both datasets ($p = 0{,}018$). Therefore, the first research hypothesis is accepted.

Mean positional error is 2.717 centimeters (HoloLens) and 3.003 centimeters (VirSat) respectively (see figure 10). Angular error averaged 3.755 degrees for the HoloLens interface and 0.588 degrees for the VirSat interface (figure 11). A test for normality revealed that both the measurements for positional and angular error are not normally distributed. Therefore, a Wilcoxon signed-rank test was used for significance testing with a critical value of 208. The test statistic for differences in positional error is above the critical value ($w = 322$). Consequently, the second research hypothesis is rejected. In contrast, for differences in angular error, statistical significance could be verified ($w = 27$). The third research hypothesis is therefore accepted.

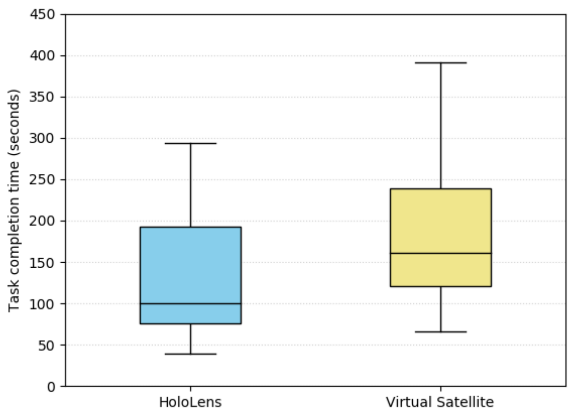

Figure 8: Box plots for completion time.

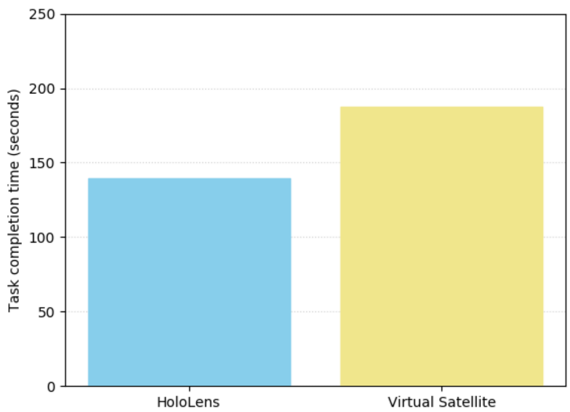

Figure 9: Mean completion time per task.

7 Discussion

Lower task completion times for the HoloLens interface arguably arise because of a variety of reasons. Most importantly, integrated control allows for efficient and coordinated manipulation of component's translation and orientation. In contrast, manipulations using the VirSat interface are more cumbersome: A user has to manipulate each parameter sequentially by clicking into a text field and editing a value. Moreover, target acquisition using the HoloLens interface is simpler. Objects can be easily selected and manipulated by looking at

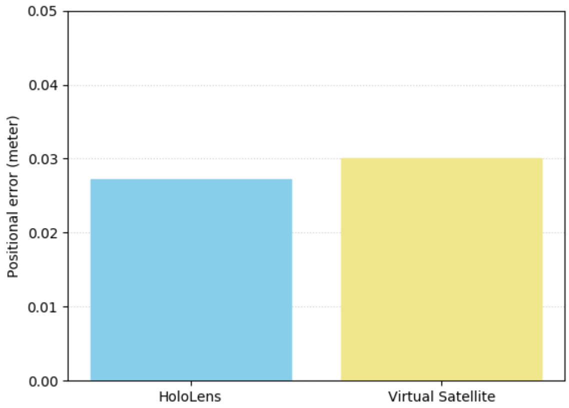

Figure 10: Mean positional error per task.

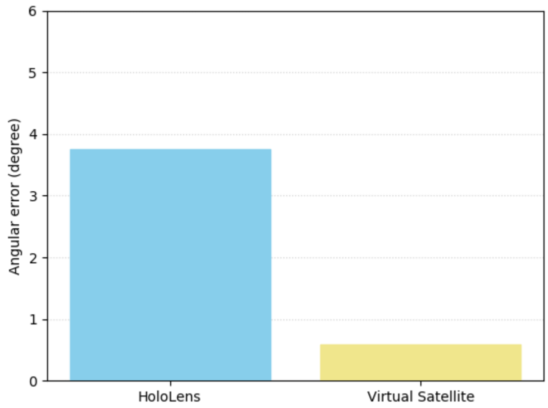

Figure 11: Mean angular error per task.

and dragging them. Using VirSat, a component has to be searched within a component hierarchy before it can be manipulated. Users have to constantly switch focus between views for parameter manipulation and 3D visualization. Often participants edited several parameters before realizing they have been editing the wrong component. However, it is important to point out that VirSat's user interface is clearly lacking a responsive 3D visualization view. These problems could easily be addressed by clicking a component in the 3D view to open up the parameter configuration view and supporting dragging by mouse.

Regarding mean positional error, statistically significant differences were not observable. This is a surprising result since VirSat's widget-based GUI provides all necessary information to simply calculate the ideal position and alignment. Still, most participants approximated a solution by testing different values for the object's translational and rotational parameters. Both approaches have proven to be time consuming, contributing to higher task completion times. These results could verify the utility of motion scaling for performing fine grained manipulations efficiently. However, the position and rotation values presented by VirSat's GUI were clearly appreciated by most participants. Numerical values were often referred to for reviewing a component's position and orientation in 3D space. This clearly indicates that a text-based evaluation of object parameters could supplement the purely visual assessment of the proposed 3D interface. In contrast, results for mean angular offset clearly favor the VirSat interface. Two main factors may have contributed to this observation. Participants could easily copy the angular parameters from the text fields of the target's neighboring cube and add a multiple of 90 degrees to align it. Therefore, with regard to orientation the posed tasks were particularly easy to solve. In contrast, participants had considerable difficulties understanding the rotation metaphor using the HoloLens interface. As expected, making purposeful rotations was especially hard. Randomly probing an object's rotation behavior often helped participants to find an accurate axis of rotation quickly. However, to achieve precise alignment with a target transformation, small adjustments about other axes usually had to be made. This would complicate making minuscule adjustments, because in again searching the correct axis of rotation new alignment errors were easily introduced to a component's rotation. A metaphor enabling more purposeful and intentional rotations seems necessary.

8 Conclusion

In this paper we presented an integrated HoloLens application for spacecraft configuration, employing a custom 3D user interface. We evaluated the 3D user interface against the VirSat interface. Based on the test results, we conclude that the proposed 3D interface is a valuable extension of the desktop-based visualization of Virtual Satellite and thus useful for spacecraft configuration. As previously mentioned, a major advantage of the proposed application is that both the 3D interface and the desktop-based interface can be used simultaneously. Nonetheless, the evaluation revealed plenty of room for improvement regarding the chosen interaction metaphors. Especially, it is shown that integrated manipulation of three rotation parameters in 3D space is not sufficient to meet the requirements. Instead, rotation based on constraints and handles will be investigated in the future. Also, seamless integration of translation and rotation within a single input technique, allowing for coordinated motion within all six degrees of freedom, should be tested against improved rotation and translation techniques in separate modes. For that matter, extending the capabilities of the HoloLens by integrating external input devices supporting finger tracking can be considered. Moreover, collaborative object manipulation has not been considered yet. This could be an important aspect given the collaborative nature of Concurrent Engineering approaches. Ultimately, further advancements in AR and VR technology can open up new possibilities for 3D interactions.

References

[BKLP04] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

[Cen18] Windows Dev Center. Gestures, March 2018. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/gestures (Accessed: 10 Jun. 2018).

[Col18] Seth Colaner. What's inside microsoft's hololens and how it works, August 2016. URL: http://www.tomshardware.com/news/microsoft-hololens-components-hpu-28nm,32546.html (Accessed 10 Jun. 2018).

[FDM+17] Philipp M. Fischer, Meenakshi Deshmukh, Volker Maiwald, Dominik Quantius, Antonio Martelo Gomez, and Andreas Gerndt. Conceptual Data Model - A Foundation for Successful Concurrent Engineering. *Concurrent Engineering Research and Applications*, November 2017.

[Fit92] Paul M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 121:262–269, 10 1992.

[FKK07] Scott Frees, G. Drew Kessler, and Edwin Kay. Prism interaction for enhancing control in immersive virtual environments. *ACM Trans. Comput.-Hum. Interact.*, 14(1), 2007.

[FWG12] Philipp M. Fischer, Robin Wolff, and Andreas Gerndt. Collaborative satellite configuration supported by interactive visualization. In *2012 IEEE Aerospace Conference*, pages 1–11, 2012.

[Mot18] Leap Motion. Home page, August 2018. URL: https://www.leapmotion.com/ (Accessed: 28 Aug. 2018).

[MRFJ16] Daniel Mendes, Filipe Relvas, Alfredo Ferreira, and Joaquim Jorge. The benefits of dof separation in mid-air 3d object manipulation. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology*, VRST '16, pages 261–268, New York, NY, USA, 2016.

[ND13] Thi Thuong Huyen Nguyen and Thierry Duval. Poster: 3-point++: A new technique for 3d manipulation of virtual objects. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 165–166, March 2013.

[Par95] Lawrence M. Parsons. Inability to reason about an object's orientation using an axis and angle of rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6):1259–1277, 1995.

[WAB93] Colin Ware, Kevin Arthur, and Kellogg S. Booth. Fish tank virtual reality. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pages 37–42, New York, NY, USA, 1993. ACM.

[ZMR97] Shumin Zhai, Paul Milgram, and Anu Rastogi. Anisotropic human performance in six degree-of-freedom tracking: an evaluation of three-dimensional display and control interfaces. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 27(4):518–528, Jul 1997.