

RESEARCH

Open Access



Rear-screen and kinesthetic vision 3D manipulator

Chao-Chung Yang*, Shih-Chung Jessy Kang, Hsiang-Wen Yang and Tzong-Hann Wu

Abstract

Background: The effective 3D manipulation, comprehension, and control of 3D objects on computers are well-established lasting problems, which include a display aspect, a control aspect, and a spatial coupling between control input and visual output aspect, which is a debatable issue. Most existing control interfaces are located in front of the display. This requires users to imagine that manipulated objects that are actually behind the display exist in front of the display.

Methods: In this research, a Rear-Screen and Kinesthetic Vision 3D Manipulator is proposed for manipulating models on laptops. In contrast to the front-screen setup of a motion controller, it tracks a user's hand motion behind screens, coupling the actual interactive space with the perceived visual space. In addition, Kinesthetic Vision provides a dynamic perspective of objects according to a user's sight, by tracking the position of their head, in order to obtain depth perception using the "motion parallax" effect.

Results: To evaluate the performance of "rear-screen interaction" and Kinesthetic Vision, an experiment was conducted to compare the front-screen setup, the rear-screen setup with Kinesthetic Vision, and the rear-screen setup without it. Subjects were asked to grasp and move a cube from a fixed starting location to a target location in each trial. There were 20 designated target locations scattered in the interactive space. The moving time and distance were recorded during experiments. In each setup, subjects were asked to go through five trial blocks, including 20 trials in each block. The results show that there are significant differences in the moving efficiency by repeated measures ANOVA.

Conclusion: The Rear-Screen and Kinesthetic Vision setup gives rise to better performance, especially in the depth direction of movements, where path length is reduced by 24%.

Keywords: 3D manipulator, Virtual reality, VR, Rear-screen, Kinesthetic vision, Eye-hand coordination, Hand-eye coordination

Background

3D computer graphics technology allows people to display 3D models on computers. As the technology advances, it has become widely used in various industries including animation, gaming, and computer-aided design. However, the limitations of display and control devices still introduce difficulties when comprehending and interacting with 3D models. Further, the spatial coupling between a perceived visual location and a manipulating location of models is still a debatable issue.

The first issue is the two-dimensional limitation of display devices. Although models are in three

dimensions, it still takes efforts to present them stereoscopically. To make models "pop out" of screens, 3D viewers commonly use the technique of presenting two offset images separately in different eyes, requiring extra head-worn devices (Eckmann 1990). Another way to enhance stereoscopic perception is by using "motion parallax" effects, which is the relative displacement of viewed models by changing observers' positions (Rogers and Graham 1979). On the other hand, Projection Augmented Model utilized a physical model, which is projected with computer images. This method present 3D models in a realistic looking. However, the pre-defined geometry

* Correspondence: ccyang@caece.net

Department of Civil Engineering, National Taiwan University, Taipei, Taiwan

shape and high precision of objects tracking and projecting is required (Raskar et al. 1998).

The second issue is the limitation of control devices. Dominant 2D input devices, which allow fine control of two-dimensional motion, are inappropriate for 3D manipulating due to the limited number of degrees-of-freedom (DoF). As a result, a mouse with virtual controllers for 3D manipulating has been discussed and evaluated in conjunction in several previous studies (Chen et al. 1988) (Khan et al. 2008). To overcome the limited DoF, controllers with three or more DoF are also developed for enhancing usability in 3D interactions (Hand 1997).

The last issue is coupling between control input and visual output spaces. Humans process visual cues received from eyes and proprioception from hands guide the movements of hands to reach and grasp models; this is called eye-hand coordination (Johansson et al. 2001). Good eye-hand coordination can reduce the mental burden during manipulation. However, most motion controllers decouple the perceived visual space (which is behind the display) and interactive space of models in front of the display (so called “front-screen” in the following chapters). Some people consider that, although this method follows the usual method of computer use, it may separate eye-hand coordination. Users’ brains need to make a semi-permanent adjustment of the spatial coupling between these spaces (Groen and Werkhoven 1998). This adaptation leads to negative after-effects of eye-hand coordination (Bedford 1989). To discuss these issues, some related works about spatial coupling problems are reviewed in the next section.

Related work

In previous research, there have been two kinds of interaction methods to solve the problem of spatial coupling.

Immersive display

Head-mounted displays (HMD) immerse users in the virtual environment. As a result, all visual perception of space is virtual, and the coupling problem no longer exists. HMD are widely used in virtual environment navigation. Newton et al. proposed the Situation Engine, which combines simulated environment with HMD and gestural control, to provide a hyper-immersive construction experience (Newton et al. 2013). However, the disadvantage is that it is relatively expensive, and it is not appropriate for extended use because it can cause dizziness and there is a need to coordinate between the virtual space and real input space (Hall 1997). Also, it focuses on large-scale 3D environment exploration rather than the manipulation of models.

Existing rear-screen interaction

Another method is to “partially” bring users into a virtual environment. The method combines Augmented Reality (AR) technologies, which fuses virtuality and reality, and the rear-screen setup, which makes users enter the fused and interactive environment visually by placing it at the back of displays. Kleindienst invented a viewing system for object manipulation, by coinciding the manipulation spaces as well as the real and virtual spaces in the viewing device (Kleindienst 2009). Holodesk, combining the optically transparent display with a Kinect camera for sensing hand motion, makes users interact with 3D graphics directly (Hilliges et al. 2012). Using the same concept, SpaceTop with the transparent OLED display is a desktop workspace that makes it easy for users to interact with floating elements on the back of the screen (Lee et al. 2013). The rear-screen idea is also brought to touch-screen devices for preventing fat-finger problems (Baudisch and Chu 2009).

In this contribution, we emulate a “rear-screen” using a laptop and a motion controller, which is not required special devices and able to be set up simply, and compare between “rear-screen” and “front-screen” tasks to validate the superiority of rear one in term of the efficiency and fatigue, due to the spatial coupling.

Method

Rear-screen and kinesthetic vision 3d manipulator

In this research, we proposed the rear-screen and kinesthetic vision 3D manipulator with a simple physical setup. Users are able to manipulate 3D models behind computer screens. Using the proposed method, the “Real Space Virtual Reality” makes the perceived virtual space and real interactive space coincident. We introduce the details of the research in this section, which is divided into the input and output modules: Rear-Screen Interaction and Kinesthetic Vision.

Rear-screen interaction

In the virtual environment, virtual simulated hands are constructed in the same dimension and position with real hands behind the screen. Users enter their hand into the virtuality and interact directly with 3D models (Fig. 1). The models in the virtuality should be constructed in the correct dimensions by referencing the scale between the virtual eye coordinates and the actual eye coordinates.

Kinesthetic vision

Positions synchronizing between virtual and actual eyes

The purpose of this part is to present the appropriate virtual scene by synchronizing the actual and virtual eye positions (Fig. 2). When the virtual and actual eyes move

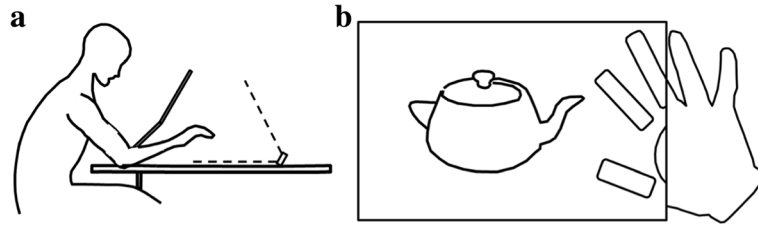


Fig. 1 Schematic of rear-screen interaction: (a) side view of the physical setup and (b) screen view

simultaneously, the relative displacement of the viewed objects, the so-called “motion parallax”, provides a visual depth cue.

$$x_V = \frac{W_V}{W_A} \cdot x_A \quad (1)$$

$$y_V = \frac{H_V}{H_A} \cdot y_A \quad (2)$$

$$z_V = \frac{D_V}{D_A} \cdot z_A \quad (3)$$

x_V, y_V, z_V is the position of the virtual eyes and x_A, y_A, z_A is the position of the actual eyes. The coordinate origin is at the center of the screen and the near plane. W_V is the width of the near plane, and W_A is the width of the screen view. H_V is the height of the near plane, and H_A is the height of the screen view. D_V is the distance from of the virtual eye coordinates origin to the near plane center, and D_A is the distance from of the actual eye coordinates origin to the screen center.

Frustum calibration

In order to simulate the shape of the actual viewing frustum through a virtual frustum, the position of the user's

eyes relative to the monitor is required. In Fig 3 r, l, t, b , and n are position parameters of the near clipping plane relative to the local eye coordination. Parameter f is the distance from the far clipping plane to the coordination in the z -direction; it is set to infinity. As the eyes move, the above parameters will be changed and need to be substituted into equation (4) of the projection matrix.

$$M = \begin{pmatrix} \frac{2n}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2n}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & -\frac{(f+n)}{f-n} & \frac{-2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad (4)$$

Implementation

The Rear-Screen Interaction and Kinesthetic Vision will be further introduced in this section by dividing into three parts: physical setup, software setup, and demonstration.

The physical setup

Three devices—a laptop, a webcam and a motion controller—are used. These have the advantage of

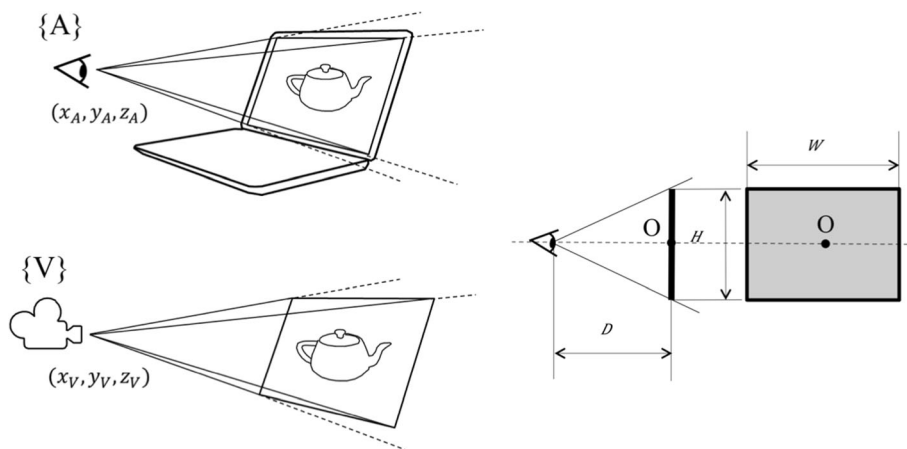
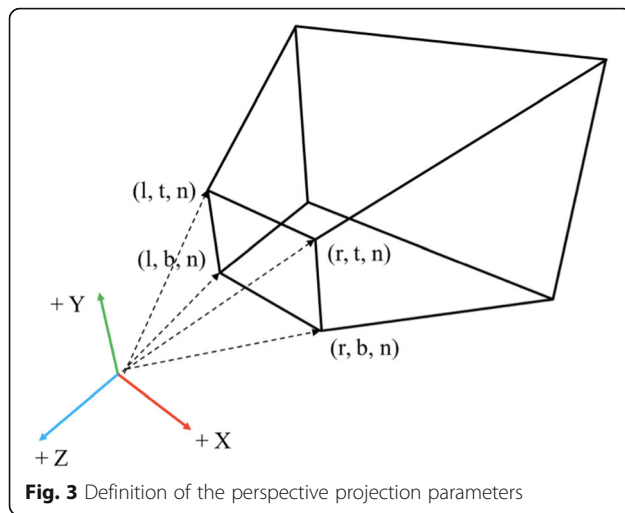


Fig. 2 Actual and virtual eye positions



being readily accessible and easy to set up. The laptop is a Lenovo x220 with 12.5" monitor, dual-core 2.3 GHz CPU and Intel HD Graphics 3000. A Logitech S5500 webcam is used for mark tracking. The webcam is set up behind users. Users are required to wear a red cap as a head tracking mark. A Leap Motion controller is a computer sensor device which detects the motions of hands, fingers and finger-like tools as input, and the Leap Motion API allow developers to obtain tracking data for further use (Fig. 4). The effective range of the controller extends from 25 to 600 mm above the device, with 0.01 mm accuracy (Leap Motion Inc 2010.)

The software setup

The Unity game engine is chosen to construct the game environment, developed in C#. OpenCV libraries are used to implement the mark tracking function, and are integrated with Leap Motion API.

System demonstration

We constructed a realistic environment similar to the real environment behind the screen, and kinesthetic vision was implemented to provide the correct perspective (Fig. 5).

Experiments and evaluation

This section will introduce the experimental method for performance evaluation, including experiment procedures, participants, and performance measurement methods.

Experiment design

We set three conditions to compare the performance of our rear-screen setup and standard setups: Rear-Screen Interaction with Kinesthetic Vision (RIK), Rear-Screen Interaction (RI), and Front-Screen Interaction (FI) (Fig. 6). By comparing RIK and RI, we attempt to ascertain if the motion parallax effect is effective for depth perception. Likewise, RI is compared with FI to confirm the superiority of rear- to front-screen in eye-hand coordination.

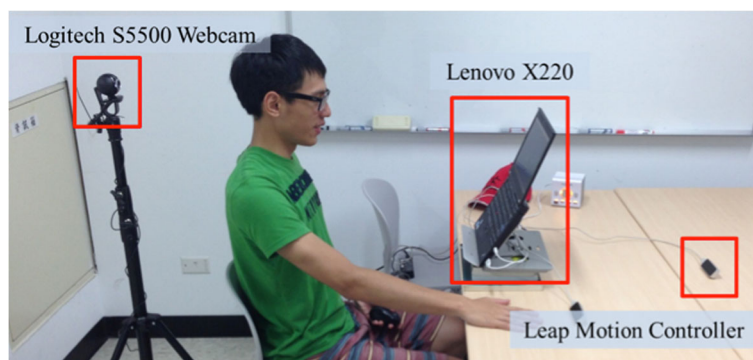
Participants

We recruited 12 participants for the experiments. All participants are male and ranged from 22 to 25 years of age. The participants are right-handed and have normal vision. They were also required to have at least 6 months' experience using software with 3D models manipulation functions, such as SketchUp, Revit, and Unity3D.

Procedures

Phase I: Introduction and Preliminary Practice

First, users are introduced the overview of the experiment, including the physical setup and the software setup. Then, participants are required to practice the grab, release and move actions. The most important aim of this section is to make the user familiar with the setup and control device, avoiding subjective factors.



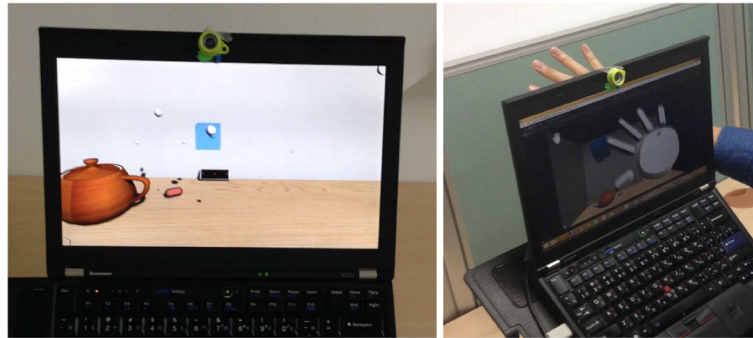


Fig. 5 Rendering results of kinesthetic vision and a simulated hand

Phase II: Formal Test: Moving Objects

Users are asked to grab and move a green cube (starting position) to a red cube (target position) in a trial (Fig. 7). The interaction depth is about 60 cm. Starting and target positions are coupled beforehand to avoid in-condition variance with random orders. Five yellow cubes appear in random positions to avoid temporary position memory.

Each user has to conduct three sets of tasks according to the three aforementioned conditions. Each set of tasks are divided into 5 blocks and each block contains 20 trials.

Phase III: Formal Test: NASA-TLX

Last, participants conduct the NASA Task Load Index (NASA-TLX) (Hart and Staveland 1988), coupled with the fatigue scale and the overall scale after each set.

Each condition takes about 30 min, including rest time between each block for fatigue prevention. After the quantitative test, we interview users about their impressions to obtain qualitative results.

Performance measurement

Zhai reported six basic aspects to the usability of a six DoF input device: speed, accuracy, ease of learning, fatigue, coordination, and device persistence and acquisition (Zhai 1998). Excluding device persistence and acquisition, which is not applicable here, we describe the method for qualitatively measuring each of the above aspects to evaluate the performance of the rear-screen kinesthetic 3D manipulator.

- **Speed:** The task completion time is divided into 2 periods: the object acquisition time and the object moving time. The measurement of the acquisition time is triggered once the virtual hand is visualized, and ends once the user grabs the object. The moving time is triggered once the user grabs an object, and ends once the object reaches the target location and the space bar is subsequently pressed.
- **Accuracy:** When the user presses the space bar, the distance between the centers of the object and the target is measured.
- **Ease of learning:** We compare the performance between blocks of trials to evaluate whether the user improves by measuring the slope of the regression line between blocks of trials.
- **Fatigue:** We reference the scaling of NASA-TLX to rate the fatigue.
- **Coordination:** The ratio between actual trajectory length and the most efficient trajectory length is measured. In our design, the most efficient trajectory is the straight-line distance between two objects. The lengths in the x, y, and z-directions are also recorded.

Result

Table 1 shows only two significant differences between setups of most of the usability aspects according to repeated measures ANOVA: Coordination ($F(2, 22) = 3.919$, $*p < 0.05$) and Grab time ($F(2, 22) = 4.157$, $*p = 0.029 <$

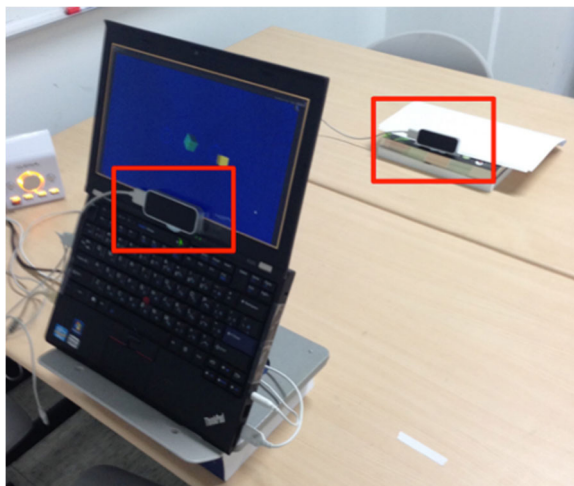


Fig. 6 Front-screen setup and rear-screen setup

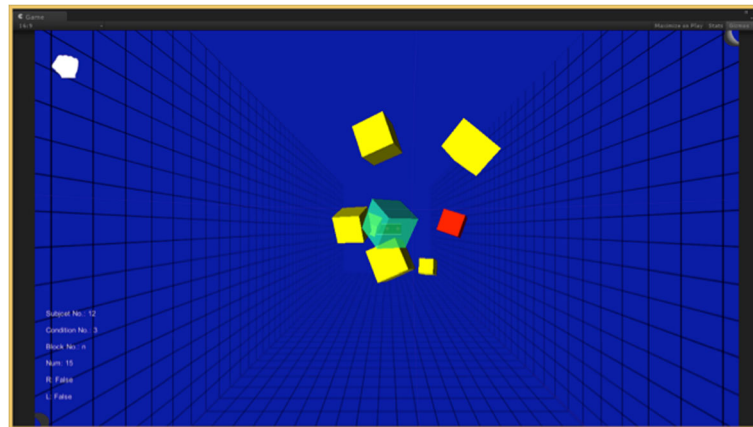


Fig. 7 Experiment software setup

0.05). When we visualize Grab time differences between FI, RI, and RIK (Fig. 8.), the figure indicates that the real significance is between FI and RI, but not RI and RIK. This matches our expectations. Under the RIK conditions, participants move left and right in order to distinguish the depth, however their hand is still outside of the screen. As a result, participants cannot distinguish the position of their hand with respect to the green box.

Figure 8. Grab time for Front-Screen Interaction (FI), Rear-Screen Interaction (RI) and Rear-Screen Interaction with Kinesthetic Vision (RIK) Variants of the rear-screen and kinesthetic vision 3D manipulator. Error bars represent \pm SEM (Standard Errors of the Mean.)

In Fig. 9a, in accordance with our expectations, RIK has a better ratio than RI, and RI also has a better ratio than FI. However, these ratios only range from 0.503 to 0.556, which does not show obvious significance. Consequently, we focus on coordination in the z-direction, according to our research goal. In Fig. 9b, the differences of coordination in the z-direction for the three conditions are highly significant according to repeated measures ANOVA ($F(2, 22) = 27.751$, $**p < 0.001$) (F -value means variation between sample

means divided by variation within the sample.) The rear-screen interaction with kinesthetic vision has the most efficient z-direction trajectory ratio (0.597), followed by one without kinesthetic vision (0.549) and the front-screen interaction (0.453). Also, post-hoc pair-wise comparisons (Bonferroni-corrected) showed significant differences between all conditions ($p < 0.05$) (p , $*p$ and $**p$ stand for different significant level from low to high.)

Figure 9a and b. Coordination ratios across the Front-Screen Interaction (FI), Rear-Screen Interaction (RI) and Rear-Screen Interaction with Kinesthetic Vision (RIK) Variants of the rear-screen and kinesthetic vision 3D manipulator: (a) Coordination in all directions; (b) Coordination in the Z-direction. (Error bars represent \pm SEM.)

Discussion

No significant difference in speed

Surprisingly, the object move time shows no significant difference between the three conditions ($p > 0.01$). We observed that movement speed varies according to personal habits.

Distraction and difficulties in eye-hand coordination

From users' feedback in the interviews, we learned users are prone to be distracted by the virtual and actual hands in the FI setup. As a result, the user finds it difficult to explore in the depth direction, leading to less efficient trajectories.

Applications

Design review

Design Review (DR) is a critical control point throughout the product development process to evaluate a design against its requirements. By combining of CAD

Table 1 Significance of usability aspects by repeated measures ANOVA

Performance		Significance
Speed	Grab Time	.029
	Moving Time	.550
Ease of Learning	Grab Time	.303
	Moving Time	.311
	Coordination	.860
Fatigue		.675
Coordination		.035

Texts in bold indicate significant differences between the observations ($*p < 0.05$)

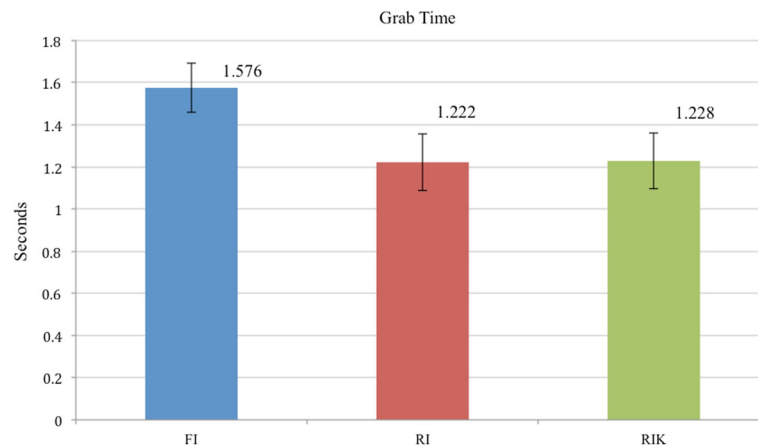


Fig. 8 Grab time for front-screen interaction (FI), rear-screen interaction (RI) and rear-screen interaction with kinesthetic vision (RIK) variants of the rear-screen and kinesthetic vision 3D manipulator. Error bars represent \pm SEM (Standard errors of the mean.)

and VR techniques, Digital or Virtual Prototyping allows to advance decisions in the early review phase to save time and cost (Bullinger et al. 2000). The review process of digital models requires several rounds of 3D manipulation in order to comprehend a design in sufficiently great detail. As the results, depth perception and eye-hand coordination are crucial for efficient exploring in a 3D virtual environment.

Gaming

Eye-hand coordination, i.e. visuomotor coordination, plays an important role in playing video or computer games (Spence and Feng 2010). Players must respond accurately and quickly to visual information. Coupling between virtual and real spaces reduces the extra effort required for spatial adaption, enhancing user experiences in gaming.

Eye-hand coordination training and testing

Taking the advantage of eye-hand coordination ability in our design, the setup is potential to be developed into training or testing tools. In the previous research, a VR-based surgical simulator is validated that it is able to differentiate between different eye-hand coordination skills (Yamaguchi et al. 2007).

Conclusion

We propose a rear-screen and kinesthetic vision 3D manipulator, which is a novel 3D object manipulation method with a simple setup. Users are allowed to interact with a virtual object directly behind the screen. The components of the rear-screen and kinesthetic vision 3D manipulator are described and implemented in this research. Finally, experiments are conducted to evaluate the design.

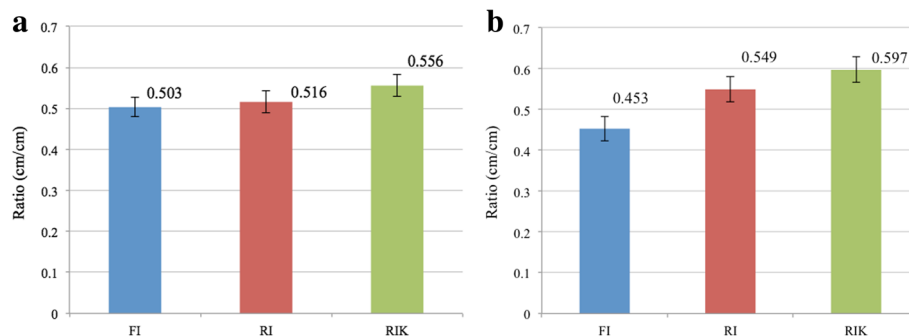


Fig. 9 Coordination ratios across the front-screen interaction (FI), rear-screen interaction (RI) and rear-screen interaction with kinesthetic vision (RIK) variants of the rear-screen and kinesthetic vision 3D manipulator: **(a)** coordination in all directions; **(b)** coordination in the Z-direction. (Error bars represent \pm SEM.)

The experimental results show there is a significant difference in coordination in the z-direction between FI, RI and RIK. Therefore, objects whose trajectory is in the depth direction are more efficiently manipulated using the rear-screen and kinesthetic vision 3D manipulator than using the standard setup. In general term, the kinesthetic sense improves users' depth perception. The finding shows the possibility and value of installing sensors for use in the design review and gaming domains.

Acknowledgements

None.

Funding

No funding to declare.

Authors' contributions

HWY, THW and CCY did the literature review and drafted the manuscript together. CCY developed the system, implemented and analyzed the validation experiment. SCK was the adviser and proof-read the article. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 June 2016 Accepted: 22 May 2017

Published online: 15 June 2017

References

- Baudisch, P., & Chu, G. (2009). *Back-of-device interaction allows creating very small touch devices* (Proceedings of the 27th international conference on human factors in computing systems - CHI 09). New York, USA: ACM Press. 1923.
- Bedford, F. L. (1989). Constraints on learning new mappings between perceptual dimensions. *J Exp Psychol: Hum Percept Perform*, *Am Psychol Assoc*, *15*(2), 232.
- Bullinger, H.-J., Warschat, J., & Fischer, D. (2000). Rapid product development — an overview. *Comp Ind*, *42*(2-3), 99–108.
- Chen, M., Mountford, S., & Sellen, A. (1988). A study in interactive 3-D rotation using 2-D control devices. *ACM SIGGRAPH Comput Graph*, *22*(4), 121–129.
- Eckmann, R. (1990). "Apparatus for assisting viewing of stereoscopic displays." US Patent No 4,925,270.
- Groen, J., and Werkhoven, P. J. (1998). "Visuomotor Adaptation to Virtual Hand Position in Interactive Virtual Environments." Presence: Teleoperators and Virtual Environments, MIT Press 238 Main St., Suite 500, Cambridge, MA 02142-1046 USA info@mit.edu, *7*(5), 429–446.
- Hall, T. W. (1997). "Hand-Eye Coordination in Desktop Virtual Reality." CAAD futures 1997, Springer, Netherlands 177–182.
- Hand, C. (1997). A survey of 3D interaction techniques. *Comput Graph forum*, *16*(5), 269–281.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): results of empirical and theoretical research. *Adv Psychol, Elsevier*, *52*, 139–183.
- Hilliges, O., Kim, D., Izadi, S., Weiss, M., & Wilson, A. (2012). *HoloDesk: direct 3D interactions with a situated See-through display* (pp. 2421–2430). Texas: Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI'12.
- Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *J Neurosci: Official J Society Neurosci*, *21*(17), 6917–6932.
- Khan, A., Mordatch, I., Fitzmaurice, G., Matejka, J., & Kurtenbach, G. (2008). *ViewCube* (Proceedings of the 2008 symposium on interactive 3D graphics and games - SIGD'08, pp. 17–25). New York, USA: ACM Press.
- Kleindienst, O. (2009). "Viewing system for the manipulation of an object." US Patent No 8,767,054.
- Leap Motion Inc. (2010). "Leap Motion." <<https://www.leapmotion.com/>> (Jun. 5, 2014).
- Lee, J., Olwal, A., Ishii, H., and Boulanger, C. (2013). "SpaceTop: integrating 2D and spatial 3D interactions in a see-through desktop environment." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'13, New York, 2–5.
- Newton, S., Lowe, R., Kember, R., & Wang, R. D. S. (2013). *The situation engine: a hyper-immersive platform for construction workplace simulation and learning*. London: the 13th International Conference on Construction Applications of Virtual Reality.
- Raskar, R., Welch, G., and Fuchs, H. (1998). "Spatially augmented reality." First IEEE Workshop on Augmented Reality (IWAR'98), San Francisco, 11–20.
- Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*(2), 125–134.
- Spence, I., & Feng, J. (2010). Video games and spatial cognition. *Rev Gen Psychol, Educ Publishing Found*, *14*(2), 92.
- Yamaguchi, S., Konishi, K., Yasunaga, T., Yoshida, D., Kinjo, N., Kobayashi, K., Ieiri, S., Okazaki, K., Nakashima, H., Tanoue, K., Maehara, Y., & Hashizume, M. (2007). Construct validity for eye-hand coordination skill on a virtual reality laparoscopic surgical simulator. *Surg Endosc*, *21*(12), 2253–2257.
- Zhai, S. (1998). User performance in relation to 3D input device design. *ACM SIGGRAPH Comput Graph*, *32*(4), 50–54.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)