Interaction in an Immersive Collaborative Virtual Reality Environment: A Comparison Between Leap Motion and HTC Controllers

Elisa Gusai, Chiara Bassano, Fabio Solari, and Manuela Chessa(✉)

Department of Informatics, Bioengineering, Robotics and System Engineering - DIBRIS, University of Genoa, Genoa, Italy
chiara.bassano@dibris.unige.it, {fabio.solari,manuela.chessa}@unige.it

Abstract. The spread of immersive virtual reality technologies, e.g. low-cost head-mounted-displays, has opened the way to the development of collaborative and interactive virtual environments, which can be exploited to obtain effective exergames. One open issue is how to obtain a natural interaction within these environments. This paper presents a prototype of collaborative environment, where users, immersed into virtual reality, can manipulate objects by using the HTC Vive controllers or the Leap Motion. We investigate which interaction modality is better by taking into account both objective measurements (e.g. the number of positioning errors) and qualitative observations.

Keywords: Human-computer interaction · VR-based exergames
HTC vive · Leap motion · Immersive VR

1 Introduction

The spread of virtual reality (VR) technologies and innovative devices has contributed to the expansion of their application domains much further than the ones that could be defined as traditional fields of interest. VR allows to isolate the user from real world and teleport her/him in a virtual environment in which the only limit is imagination. Immersivity and an intuitive interface have certainly contributed to increase the interest in this computer technology, so, even if virtual reality is still a prerogative in sectors such as military industry and simulation, it is starting to play an important role also in the game and entertainment sectors. Moreover, researchers have understood its huge potential and are investigating the use of VR in many others different fields: from the design of serious games, which combine entertainment with educational purpose [1], to support surgeons in diagnosis, operation planning and minimal invasive surgery and in rehabilitation contexts [2,3]; from products design, assembly and prototyping process to cultural heritage applications such as virtual museums [4] or historic sites modeling [5,6]. It is now evident how virtual reality is becoming a significant part of consumer everyday life. Thinking of VR mainly as a way

S. Battiato et al. (Eds.): ICIAP 2017 International Workshops, LNCS 10590, pp. 290–300, 2017.
https://doi.org/10.1007/978-3-319-70742-6_27

to isolate people from real world and provide an individual experience, however, would be wrong. Indeed, another sector that has particularly benefited of VR is certainly the one of *Collaborative Virtual Environments* (CVE). Several studies have proved that, when multiple subjects who have to carry out a common task share the same work space, this cooperation can bring a series of great advantages [7,8]. While the concept of CVE is quite clear, the actual creation of this cooperative work space remains an open problem and a research topic.

This paper mainly focuses on the interaction between user and virtual objects and our contribution consists in the evaluation of the intuitiveness and naturalness of the human-computer interaction. We have designed a simple exergame in which the player is asked to grab and move objects. We use the HTC Vive Head-Mounted Display (HMD) to visualize the scene and two different modalities to interact with it: one modality uses the controllers that HTC Vive itself provides; the other modality uses the Leap Motion, a low-cost hand tracker meant to provide a natural interaction in VR. Our aim is to identify which one between the two devices represents the best solution to be employed within the field of a manipulation task, both in terms of performances in accomplishing the task and in terms of the quality of the experience of the involved subjects.

2 Related Works

Recently, several studies took in examination different aspects of CVEs in several domains of application. One of the greatest results has been obtained by the comparison of the CAVE system's performances and the HMD's ones in a collaborative task of abstract data visualization [9]. Three different aspects have been taken into account: functionality, esteem of collaboration degree and evaluation of the quality of the experience lived by the users. Under all these points of view it has came out that the employment of a low cost technology like HMD, can provide results as accurate as those obtainable through a system such us the CAVE, also offering the advantage of being more versatile and easy to handle. Beyond the purely technical aspects, when studying CVE, researchers have to take into account another important factor related to the type and degree of collaboration between the users that the system is able to support [10]. When speaking of CVEs, it is necessary to take into account the role [11] of each user in the task, as well as the factors influencing the behavior of the subject in terms of what he is allowed to do or not, and consequently the strategy he adopted in order to pursue the final aim. Of course, this has an impact also on the task's structure and on the type of reciprocal interactions the participants can establish. In fact the situation is quite different if all the participants are equally free to act or have some restrictions. In a similar context it is indispensable to preliminary make a point on the users, especially in terms of their cultural background and degree of expertise toward the task they have to perform. This particular aspect associated with CVE can be dealt through a particular approach, drawing a parallel with the interactions between users and those that rule the dynamics human-robot [12]. From that point of view it is possible to define a kind of interactive hierarchy which starts from the so called *tele-operation* strategy, in which the supervisor has to completely control the machine (or another user's actions, in case of human - human interactions). The autonomy's degree gets higher and higher as we go up from the bottom to the top of the chain of command; each one of them corresponds to a different level of *tele-assistance* strategies and results in a totally autonomous modality of work [13]. To choose the best interactive strategy to adopt it is necessary to take into account the precise task to carry out and the specific required performances. The particular type of approach employed, in fact, influences the results of the work: if in a certain context the priority is to complete the task as quickly as possible despite the precision, the best choice might be reducing supervision of the expert-user on his partner's actions; on the contrary, when accuracy is preferred with respect to the execution speed, a greater control over the actions of the beginner might require, therefore giving him a less autonomy. Finally, it is also important to define a series of customized metrics in order to effectively measure all the parameters necessary to evaluate the performances and the quality of users' experience inside VR [14,15].

3 Materials and Methods

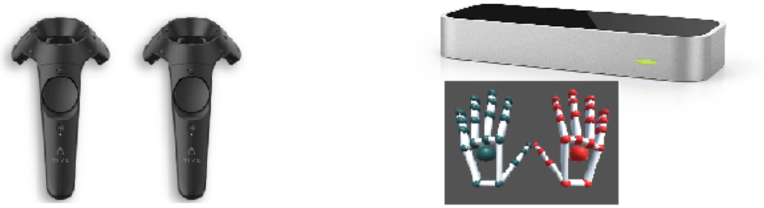
3.1 Hardware Components

The collaborative task we have defined requires the use of the following devices: a Head-Mounted Display, connected to a computer (running a 64 bit Windows10 operating system, with an Intel(R) Core(TM) i7 2.67 GHz processor, 8 GB RAM and a NVIDIA GeForce GTX680 graphic card), a pair of controllers designed for interactive VR or the Leap Motion[1]. The HMD we used is the HTC Vive[2] developed by HTC and Valve Corporation; it belongs to the class of *'room scale'* virtual reality technology. The HTC headset includes two wireless controllers and two 'lighthouse' basestations, that are able to track the head-mounted display and the controllers position in a certain area, 4.6 by 4.6 m, defined through a calibration process. The HMD consists of two screens, one per eye, with a resolution of 1080×1200 and provides a refresh rate of 90 Hz and a field of view of about 110°.

Manipulation and interaction with the virtual objects in the scene was provided by two different devices: in particular, we carried out two distinct series of experiments in which the users were able to interact with the scene using the HTC's controllers and the Leap Motion respectively. HTC Vive supplies two wireless and ergonomic controllers that allows an easy handle with just a hand (Fig. 1(a)). They have a great number of passive sensors that guarantees an extremely accurate tracking contributing to obtain an incredibly stable system. The Leap Motion (Fig. 1(b)), instead, is a hand tracker: it has been fixed on the HTC Vive with a special support placed in the frontal part of the HMD at the center. Thanks to this particular configuration it is possible to easily track and represent hands in VR. It is worth noting that in our experiments we use only

[1] https://www.leapmotion.com/.
[2] https://www.vive.com/eu/.

one HTC controller. This does not affect performances, since to accomplish the task only one controller is necessary.



(a) (b)

Fig. 1. The input devices used to interact with objects in the immersive VR environment: (a) the HTC Vive controllers and (b) the Leap Motion controller.

3.2 Software Components

The immersive VR has been developed mostly by using *Unity 3D* 5.5.1[3]. Concerning the virtual objects, we employed Blender 2.78[4], an open source computer-graphics software widely used for creating and manipulating 3D objects. Before exporting the *fbx* models obtained from Blender on Unity, it was necessary to use *Autodesk Netfabb* 2017.3[5]. Finally we downloaded some already done 3D models from *SketchUp*[6].

3.3 Subjects

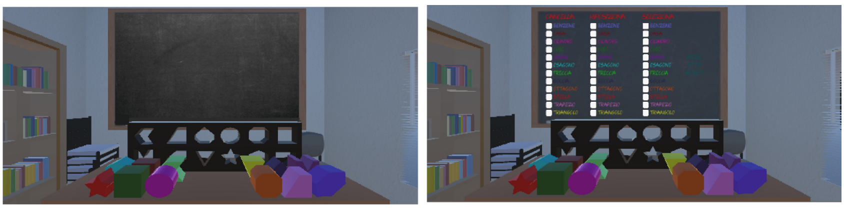
30 people, 15 males and 15 females, aged between 20 and 49 years (mean 27 $\pm$ 6.6), took part in this experiment and constitute the control group, composed by healthy subjects. Only few of the participants had already had past experiences in VR, while the majority of them had never got in touch with it, even if they had heard about it. Thereafter their experience or their confidence in virtual reality is very different. All the subjects took part in the study voluntarily and nobody perceived any kind of reward.

3.4 Collaborative Task

Since our aim is to test how people interact in an immersive VR system, in order to understand which solution is preferred by the users, we propose a very simple collaborative task: this way the obtained performances actually reflect the interaction mode rather than the complexity of the task itself. It requires the

[3] https://unity3d.com/.
[4] https://www.blender.org/.
[5] https://www.autodesk.com/products/netfabb/overview.
[6] http://www.sketchup.com/it.

presence of two users who play extremely different roles: one of them, the *main user*, has to place 12 three-dimensional objects, different in color and shape, inside the corresponding holes (Fig. 2(a)), while the other one, the *supervisor*, supports his companion's actions and helps her/him to achieve the requested target. The task structure is completely asymmetrical because the two users work through different interfaces. The *main user* employs the HMD and acts inside an extremely immersive virtual environment. In order to guarantee safety conditions during the task accomplishment, the HMD users had to play in a restricted area of the lab, corresponding to the HTC Vive calibration area. He/she can interact with virtual objects using two different modalities: HTC controllers and Leap Motion. In the first case, items are grabbed just pushing a button on the controller and dropped when the button is released; in the second case, the player can interact with objects using his own hand and natural reaching, grasping and releasing gestures. The *supervisor*, instead, uses a simple desktop application (Fig. 2(b)), she/he is able to see *main user*'s actions from a fixed point of view on the scene and can help her/him replacing single objects thrown out of the game area in their original positions, deleting some of them in order to simplify the individualization and positioning of a piece, or selecting a hole to indicate the position into which a certain piece has to be insert. The *main user* is not able to see *supervisor* user-interface, so she/he can just see a clean blackboard as shown in the images below.



(a) (b)

Fig. 2. (a) *Main user* point of view during the execution of the task. (b) *Supervisor* point of view during the execution of the task.

Regarding the application structure, it consists of three different scenes: a start menu, in which it is possible to insert the user's ID, and to select the interaction mode; a demo scene consisting in a virtual office room, where the player is free to move, explore the game area and to interact with objects according to the selected interaction mode; the main scenario, similar to the demo, in which, on the desk, there are 12 objects and a base with holes in which to put them. A blue cage, visible in the second and first scene, mark the boundaries of the calibrated game area and defines a space inside which the user can move safely (Fig. 3).

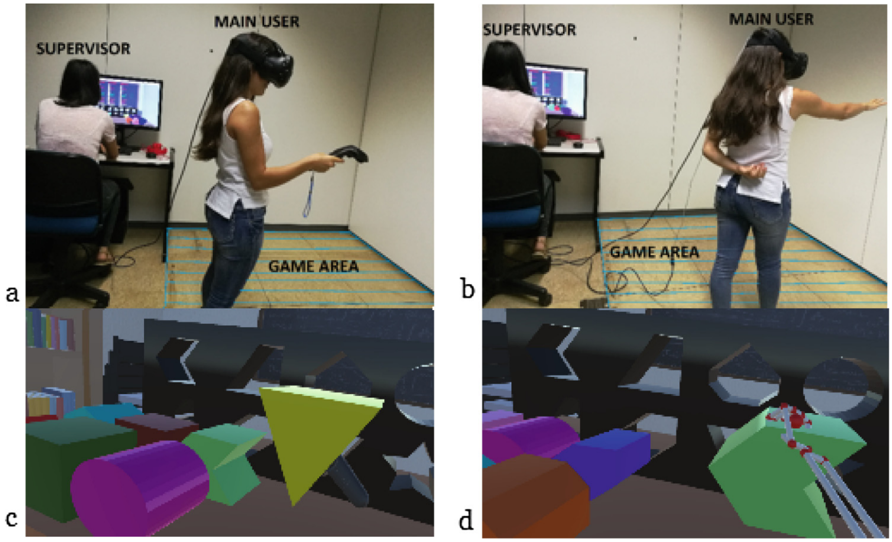


Fig. 3. Experiment setup. The *main user* wears the HTC Vive and can freely and safely move in the game area and interact with the virtual objects using the controller (a) o the Leap Motion (b). The *supervisor* sits in front of the desktop and observes main user's actions ready to intervene in case of need. Interaction with the virtual objects using the HTC controller (c) and the Leap Motion (d).

3.5 Experimental Procedure

The experiment has been carried out with 30 subjects as *main user*. The *supervisor* was always the same person. The experimental setup included two trials for each volunteer. Firstly, we asked the players to submit the Simulator Sickness Questionnaire (SSQ) [16], a 16 questions' questionnaire specifically used in the literature to evaluate the user physical status before the exposure to the virtual environment (pre-exposure). Then the users had to perform the task according to the selected interaction mode. The task ended when all the 12 objects were correctly positioned. After completing the first trial, the subjects had to submit the post-exposure SSQ; those answers have been used to evaluate users' status following the first exposure and preceding the second exposure to the VR. The second trial was executed with the interaction device complementary to the one used during the previous task. Subsequently, the users had to submit the post-exposure SSQ in order to evaluate their physical conditions at the end of the experiment. Finally, all the volunteers were asked which interaction mode they preferred. To reduce the statistical variability half of the participants carried out the first task using the Leap Motion and half using the controllers. To avoid learning the position of all the objects on the desk has been modified in the two trials.

3.6 Metrics

To evaluate the performances with the two different devices we have analyzed two distinct classes of parameters: the number of errors and the times necessary to correctly position the virtual objects. We defined a score for every task, assigning one point to each object correctly inserted (the baseline score starts from 12 points) and subtracting a certain number of points for each error committed. We associated a different coefficient to each kind of error according to their severity. The final score (S) is the result of the difference between the number of objects correctly positioned (12) and the sum of the number of repositioned (RP) and deleted (DP) piecies, and of selections (SP) and reset (R) actions, each one multiplied for its own coefficient ($c_1 \dots c_4$, empirically evaluated), as shown in the following formula:

$$S = 12 - (RP * c_1 + DP * c_2 + SP * c_3 + R * c_4)$$

Through the time analysis, we computed the total time necessary to complete the task and the average time required to position each single object.

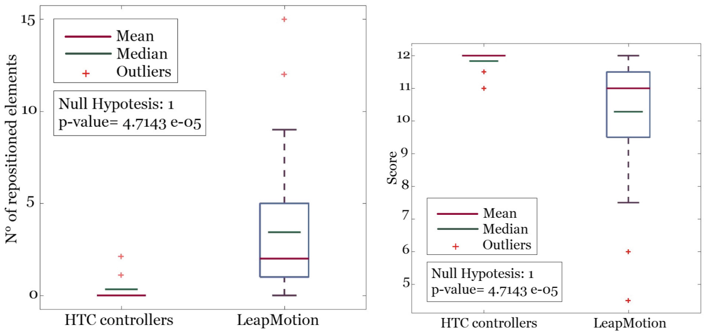
Furthermore, using the *Simulation Sickness Questionnaires* we were able to analyze the participants' *status* before and after each exposure to virtual reality in order to assess if the experience inside a virtual environment, even if brief, had physical effects on the users or not.

4 Results

The metrics defined to evaluate the task performances have been used to obtain an analysis on three levels: firstly we compared the performances obtained with the HTC controllers and with the Leap Motion; then we made a general comparison between the results relative to the first and the second trial, without considering the employed interaction device; in the end, we made a 'crossed' comparison to link the data relative to the first and the second trial taking into account the interaction mode used in both cases. Note that in all the experiments the only kind of help that the *main users* needed was the repositioning of objects; therefore, we decided to make an independent analysis on the number of repositioned elements in addition to those related to the score, the completion time and the single object's average time. In each analysis the statistical significance of the differences between data was estimate by making a *t-test* analysis.

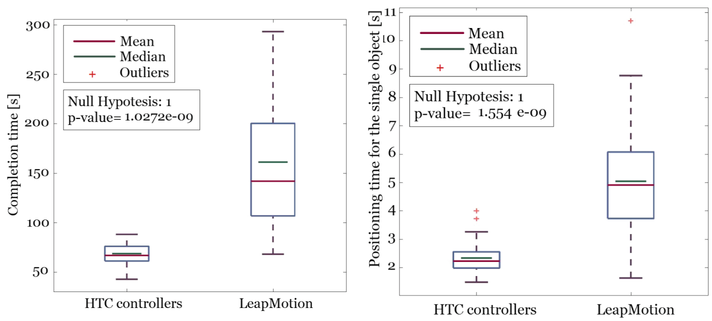
The comparison between the performances of the two interaction devices highlights that all the considered parameters are significantly lower for the HTC controller. The median value related to the number of repositioned elements is equal to 0 in case of the HTC controller and to 2 in case of the Leap Motion, while the average value is 0.3 versus 3.4 (Fig. 4(a)). The score varies from a mean of 11.8 points in the HTC controller trials to 10.3 in the Leap Motion trials (Fig. 4(b)).

The average completion time is 68.5 s for the controller and it is 92.5 s lower compared to the one obtained with the Leap Motion, which is 161.1 s (Fig. 5(a)).



(a) (b)

Fig. 4. Mean, median and standard deviation of (a) repositioned elements and (b) the final score for HTC controller and Leap Motion trials. p-value shows statistical significance.



(a) (b)

Fig. 5. Mean, median and standard deviation of (a) completion times and (b) positioning times for the single object for HTC controller and Leap Motion trials. p-value shows statistical significance.

Similarly the average positioning times for the single object are equal to 2.3 s and 5.0 s for the HTC controller and the Leap Motion, respectively (Fig. 5(b)). All these results are statistically significant.

The analysis performed on the data relative to the first and second trials shows no statistically significant differences for any of the considered parameters. Those results seem to indicate that no learning process happened between two consecutive trials. Nevertheless, this can be explained taking into account that the two series of trial have been execute with distinct interaction mode; therefore the users had to learn how to use each kind of device. The crossed comparison which considers both the order and the modality of the execution, underlines that there is not a learning pattern between first and second trials (Table 1). But comparing the performances of the first and second tasks performed using the same interaction mode, all the considered variables have a little lower value in the second trial case for both devices, though the differences are not statistically significant as explained before. The crossed analysis shows the better performances related to the first trial with the HTC controller. Similarly, the results relative to the second trial are significantly lower for the HTC controller. Finally, the SSQ analysis highlights no difference between the data acquired before and after the task execution thus indicating that the virtual environment does not cause sickness to the user. The open discussion shows how 73% of the users has preferred the HTC controller and only the 13% of them the Leap Motion device.

Table 1. Cross comparison of mean and standard deviation for each considered parameter between HTC controllers and Leap Motion relative to the first and second tasks.

Parameters	Trials	HTC controllers	Leap motion
Replaced elements	1st	0.4 ± 0.8	4.1 ± 3.7
	2nd	0.3 ± 0.6	2.8 ± 3.9
Completion time [s]	1st	73.9 ± 10.4	180.7 ± 67.9
	2nd	63.7 ± 10.9	141.1 ± 65.8
Average time per object [s]	1st	1.2 ± 0.7	2.7 ± 3.3
	2nd	1.3 ± 1.2	2.3 ± 65.8
Score/12	1st	11.8 ± 0.4	10 ± 1.8
	2nd	11.9 ± 0.3	10.6 ± 1.9

5 Conclusion

In this study, we specifically address human-computer interaction in a collaborative virtual environment, in particular comparing the Leap Motion and the HTC controller. We have developed an asymmetrical collaborative task in which one of the two involved users works inside a highly immersive 3D virtual environment and is supported in achieving his/her task by a partner who interacts with the virtual scene through a pc. The research purpose is to assess how the task's performances and the quality of the users experience inside the CVE vary according to the specific interaction device employed: the HTC Vive controller or the Leap Motion. The data analysis has included the evaluation of specific parameters related to the particular task, the measures of the users' physical conditions before and after the exposure to the virtual reality using the *Simulator Sickness Questionnaire* and the assessment of the participants quality experience inside the CVE based on the answers that they supplied during the open debate at the end of the experiment. The results we obtained highlight a statistically significant difference between the performances related to the two devices in favor to the HTC controller, in term of the number of repositioned elements, task completion time, single object average time and score. Comparing the first and second trial no relevant differences has arose, which means that the user has to learn from time to time how to interact with virtual objects through different devices. Nevertheless, the comparison between first and second trial carried out through the same device underlines that in the second case performances increase. The SSQ analysis shows that the VR exposure has not significant effects on the physical conditions of the users: this might be due also to the short time the task required to be completed. Based on the debated answers we state that the majority of the participants has been more comfortable using the HTC controller, probably because of the greater simplicity of interaction with the objects and the better stability of the system. This results was confirmed by quantitative analysis. Only 13% of the users has preferred working with the Leap Motion because of the greater natural interaction with the objects, despite obtaining worse performances; the remaining 13% liked both HTC controller and Leap Motion.

In conclusion we can declare that the system HMD - Controller offers the best performances thanks to its greater stability and accuracy; those characteristics guarantee a simpler handling providing a better experience for the user. The Leap Motion allows the user to interact with the objects in a more natural way compared to the controller, but it fails on stability and accuracy and its performances are extremely variables from subject to subject. In the future we are planning to fix some of the stability and accuracy problems related to the use of the Leap Motion and make further investigations on the manipulation of the objects in VR with HTC controllers and Leap Motion.

References

1. Monahan, T., McArdle, G., Bertolotto, M.: Virtual Reality for Collaborative e-Learning. Elsiever (2006)
2. Bouchard, S., Renaud, P., Robillard, G., St-Jacques, J.: Applications of virtual reality in clinical psychology: illustrations with the treatment of anxiety disorders. In: IEEE International Workshop 2002 HAVE Haptic Virtual Environments and Their Applications (2002)
3. Mclay, R.N., Wood, D.P., Murphy, J., Wiederhold, B.K.: A randomized, controlled trial of virtual reality-graded exposure therapy for post-traumatic stress disorder in active duty service members with combat-related post-traumatic stress disorder. Cyberpsychology Behav. Soc. Netw. **14**(4), 223–229 (2011)
4. Chessa, M., Garibotti, M., Rossi, V., Novellino, A., Solari, F.: A virtual holographic display case for museum installations. In: 2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN), pp. 69–73. IEEE (2015)
5. Li, N., Nittala, A.S., Sharlin, E., Sousa, M.C.: Shvil: collaborative augmented reality land navigation. In: Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems, pp. 1291–1296. ACM (2014)
6. Shia, Y., Dub, J., Lavyc, S., Zhaod, D.: A multiuser shared virtual environment for facility management. In: International Conference on Sustainable Design, Engineering and Construction (2016)

7. Churchill, E.F., Snowdon, D.: Collaborative virtual environments: an introductory review of issues and systems. Virtual Reality **3**(1), 3–15 (1998)
8. Okada, K.: Collaboration support in the information sharing space. IPSJ Mag. **48**(2), 123–125 (2007)
9. Cordeil, M., Dwyer, T., Klein, K., Laha, B., Mariott, K., Thomas, B.H.: Investigation of interactive strategies used in undertaking collaborative tasks. IEEE Trans. Vis. Comput. Graph. **23**, 441–450 (2017)
10. Sarmeinto, W.J., Maciel, A., Nedel, L., Collazos, C.A.: Measuring the collaboration degree in immersive 3D collaborative virtual environments. In: IEEE VR International Workshop on Collaborative Virtual Environment (3D CVE) (2014)
11. Claude, G., Gouranton, V., Arnaldi, B.: Roles in collaborative virtual environments for training. In: IEEE VR International Workshop on Collaborative Virtual Environment (3D CVE) (2017)
12. Lo, K., Zeng, T., Hu, Y.: Investigation of interactive strategies used in undertaking collaborative tasks. Appl. Sci. **7**(4), 318 (2017)
13. Sheridan, T.B.: Telerobotics, Automation, and Human Supervisory Control. MIT Press, Cambridge (1992)
14. Erfanian, A., Hu, Y.: Multi-user efficacy of collaborative virtual environments. In: IEEE 20th International Conference on Computer Supported Cooperative Work in Design (2016)
15. Geszten, D., Hmornik, B.P., Hercegfi, K.: User experience in a collaborative 3D virtual environment. In: IEEE 6th International Conference on Cognitive Infocommunications (2015)
16. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enanched method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**, 203–220 (1993)