A Study on the Role of Feedback and Interface Modalities for Natural Interaction in Virtual Reality Environments

Chiara Bassano, Manuela Chessa and Fabio Solari
Dept. of Informatics, Bioengineering, Robotics, and Systems Engineering, Italy

Keywords: Natural Human Computer Interaction, Touchful Interaction, Touchless Interaction, User Evaluation, Ecological Interface.

Abstract: This paper investigates how people interact in immersive virtual reality environments, during selection and manipulation tasks in different conditions. We take into consideration two different task complexities, two interaction modalities (i.e. HTC Vive Controller and Leap Motion) and three feedback provided to the user (i.e. none, audio and visual) with the aim of understanding their influence on performances and preferences. Although adding feedback to the touchless interface may help users to overcome instability problems, providing information about the objects state, i.e. grabbed or released, they do not substantially improve performances. Moreover, both touchful and touchless modalities have been shown to be effective for interaction. The analysis presented in this paper may play a role in the design of natural and ecological interfaces, especially in the case non-invasive devices are needed.

1 INTRODUCTION

Virtual Reality (VR) is a widespread technology and the release of new low cost devices in the last decade has certainly contribute to its success. One of the main feature of VR is the possibility of interacting with the virtual environment and the virtual objects, thus improving user involvement. However, interaction is still an issue, especially if we consider Natural Human Computer Interfaces (NHCI), i.e. interactive frameworks integrating human language and behavior into technological applications easy to use, intuitive and non-intrusive.

In this paper, we aim to analyze how people interact within VR environments, by taking into consideration two different task complexities, two interaction modalities (i.e. HTC Vive Controller and Leap Motion) and three kinds of feedback provided to the user (i.e. none, audio and visual). To this aim, we devised three experiments in two different scenarios. In the first scenario, used for Experiment 1 and taken from (Gusai et al., 2017), users have to solve a simple shape sorter game for babies with 12 shapes; while in the second scenario, used for Experiment 2 and 3, users have to assembly the Ironman movie character suit starting from building blocks.

In Experiment 1, we take into consideration a touchless interface, the Leap Motion, and we compare the effect of different feedback on performances and preferences. Similarly, in Experiment 2 we investigate if results obtained with the simple task can be extended to a more complex task. Finally, we devise an in-the-wild experiment, in order to understand how people interact in the complex scenario with state of the art technologies, HTC Vive controllers, thus obtaining a baseline for the analysis of performance in Experiment 2.

1.1 State of the Art

The level of naturalism of User Interfaces (UIs) is defined as the degree with which actions performed to accomplish the task using a certain interface correspond to the actions used for the same task in the real world (McMahan, 2011) and depends on many different factors (Bowman et al., 2012). Bare hands and human body interactions are often considered as a natural and ecological form of Human-Computer Interaction (HCI), as they are designed in order to reuse existing skills (George and Blake, 2010) through intuitive gestures requiring a little cognitive effort. However, the lack of haptic or multimodal feedback and physical constraints, the limited input information, i.e. user's head and, in some cases, hands pose, and an inaccurate tracking often restrict users to a coarse manipulation of objects (Mine et al., 1997). Vision-

Bassano, C., Chessa, M. and Solari, F.
A Study on the Role of Feedback and Interface Modalities for Natural Interaction in Virtual Reality Environments.
DOI: 10.5220/0008963601540161
In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020) - Volume 2: HUCAPP*, pages 154-161
ISBN: 978-989-758-402-2; ISSN: 2184-4321

based methods are the most widely diffuse solutions, including more or less expensive motion capture systems (e.g., Vicon system), tracking systems provided with head-mounted displays (HMDs), the Kinect for body tracking and the Leap Motion for hand tracking. These solutions present interaction and tracking challenges: they provide good tracking but in a limited area and they are prone to occlusion, noisy reconstruction and artifacts (Argelaguet and Andujar, 2013). (Augstein et al., 2017) classify input devices, considering the level of touch involved and distiguish three classes: *touchful*, e.g. mouse, controllers or touchscreens, which require the application of a physical pressure to a surface; *touchless*, e.g. Leap Motion or cameras; *semi-touchless*, combining characteristics of the two previous one. In particular, touchless input systems provide a wide range of input options, can replicate all the necessary DOFs, retrieve information while keeping sterility and allow interaction through intuitive commands and gesture. Nevertheless, they cannot provide haptic or force feedback. Research on this field has been very active in the last decades. The goal is investigating the optimal interaction method to be used for task-specific scenarios, taking into account both quantitative parameters, as a measure of performance, and subjective questionnaire, to investigate preferences: while participants perform better with the touchful or semitouchless input techniques, they prefer using the touchless one, in terms of immersion, comfort, intuitiveness, low fatigue, ease of learning and use (Augstein et al., 2017; Zhang et al., 2017). While in (Gusai et al., 2017), Leap Motion (touchless) and HTC vive controllers (touchful) were used to complete a simple shape sorter task in immersive VR. Participants performed better with the controller and preferred it for interaction, for its stability, easiness of control and predictability.

Since interfaces are task dependent, when talking about interaction modality, it is important to consider both the purpose of the interaction and the size of the space where interaction takes place. (Frohlich et al., 2006) identify four tasks categories: (i) *navigation* and *travel*, which consist in moving the viewport or the avatar through the environment; (ii) *selection*, i.e. touching or pointing something; (iii) *manipulation*, corresponding to modifying objects position, orientation, scale or shape; (iv) *system control*, the generation of an event or command for functional interaction. Our research mainly concentrates on selection and manipulation in the peripersonal or near action space (i.e. the space covered performing some steps). In this case, a VR system allowing room scale setups and a real-world metaphor for interaction would be sufficient (Cuervo, 2017).

1.2 Research Questions

In this paper, we devised 3 different experiments in immersive VR environments. Experiment 1 extends (Gusai et al., 2017)'s work by the introduction of visual and audio feedback in the Leap Motion case, with the aim of facilitating interaction. The shape sorter task proposed voluntarily requires little cognitive effort in order to focus on the role of feedback on performances. In Experiment 2, we keep the interaction modality of Experiment 1 invariant but modulate the complexity of the task, introducing a more cognitively complex assignment, the assembly of an Ironman suit. Finally, in order to distinguish between the effect of task complexity and interface used on performances and preferences in Experiment 2, we decided to replicate the complex task with a state of the art technology for VR interaction, i.e. the HTC Vive controllers. We devised a series of in-the-wild experiments with the Ironman game, using HTC Vive controllers for interaction. As in (Gusai et al., 2017), no feedback are provided.

The aim of this work is answering the following research questions.

- Can different non-haptic feedback improve performances in the touchless interface case, overcoming the lack of touch and forces provided to the user?
- Can simple feedback effectively substitute haptic feedback?
- Is there any correlation between performances and preferences on the feedback modality?
- Does task complexity influence the need for feedback when using touchless interfaces?

2 MATERIAL AND METHODS

All three experiments used the same hardware and software platform and the two scenarios were both developed in Unity 3D. The experimental setup (Fig. 1a and 2a). was composed of the HTC Vive for visualization, the Leap Motion (Experiment 1 and 2) and the HTC Vive controllers (Experiment 3). Data recorded are used to run-time create a virtual representation of the hands: objects can be grabbed and release by closing and opening the hand. HTC Vive controllers, instead, are handleable devices. In this case, items grab and release actions were performed by pressing and releasing the trigger button. With both devices, a real-world metaphor paradigm was implemented and objects could be carried around by simply moving the free hand or the hand with the controller. As stated before, (Gusai et al., 2017) have proven that controllers allow obtaining better performances, ensuring stability, ease of use and predictability. Moreover, they are preferred for interaction. On the other hand, people appreciate the idea of interacting with objects using their bare hands but often report tracking stability problems with Leap Motion, due to fast movements, occluded hand poses or by the sensor's limited field of view. In general, it is even more difficult for naive users understanding and predicting unexpected behaviour during interaction. Thus, for Experiment 1 and 2, we decided to provide feedback showing the state of the grabbed object. We implemented and compare four different conditions.

- *None*: no feedback are provided.
- Visual feedback *Visual 1*: the algorithm, taken from (Bassano et al., 2018), checks if the hand is colliding with one of the interactable shapes. When a collision is detected, the object turns gray (Fig. 1c) or the "smoothness" parameter of its material is increased (Fig. 2d). The hand's grab angle is then checked and, if it is greater than a threshold experimentally defined, the object turns violet (Fig. 1d) or the "smoothness" parameter is further increased (Fig. 2e), is attached to the palm and starts moving.
- Visual feedback *Visual 2*: shapes change color (Experiment 1) or "smoothness" (Experiment 2) once, they turn violet when grabbed (Fig. 1d).
- *Audio* feedback: when an object is grabbed a music is played, when the grabbing action finishes music is paused. This feedback modality represents the audio counterpart for *Visual 2* feedback.

We decided to implement both visual and audio feedback because in VR audio is often a minor input: users receive lots of visual information, causing delay and loss in their processing, while audio channel is almost unused. Nonetheless, in a real situation, haptic feedback is a non-visual feedback.

As our goal is understanding the variation of performances and preferences in the different experimental setups, during trials we acquired both quantitative and qualitative measurements.

As an objective meter for performance analysis, in all experiments we recorded total execution time, i.e. the time required to complete the task, and partial time, namely the time required to correctly position each object from the first grab. While partial time mainly depends on to the interaction between user and objects, total completion time can be considered as an overall evaluation, taking into account both the efficiency of the interaction, the strategy used and the time necessary to understand how to correctly position items. As a qualitative self-assessed evaluation, instead, we administered the User Experience Questionnaire (UEQ), after each trial of Experiment 1 and 2, and an overview questionnaire, after each trial of Experiment 2 and 3.

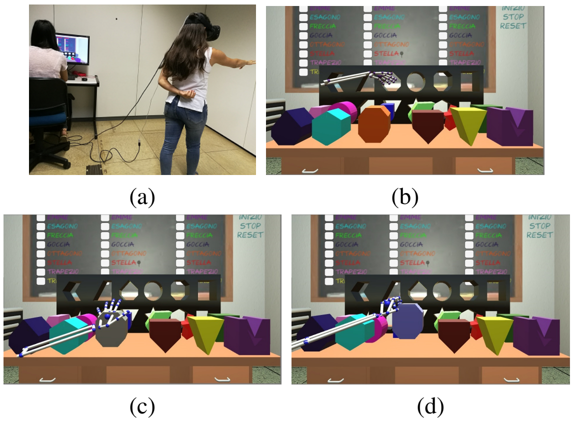


Figure 1: (a) Experiment 1 setup. (b) No collision is detected. (c) A collision between hand and object is detected. (d) Hand is closed in a grasping pose.

UEQ is a validated questionnaire for the comparison of the usability of different devices (Laugwitz et al., 2008; Schrepp et al., 2014) or, as in this case, of different feedback modalities. The questionnaire is composed of 26 questions in the form of a semantic differential and adopts a seven-stage scale for rating. Items are clustered in 6 groups: *Attractiveness* of the product; *Perspicuity*, intended as how easy is to learn to use the product; *Efficiency* of the interaction; *Dependability*, i.e. the level of control user felt; *Stimulation*, namely users' excitement and motivation; *Novelty*, in terms of innovation and creativity.

The overview questionnaire, instead, is composed of 5 questions rated in a 5-points Likert scale, where rate 1 corresponds to "Strongly disagree" and 5 to "Strongly agree": (i) How intuitive were the game commands?; (ii) How difficult was understanding to what the single suit pieces corresponded?; (iii) How difficult was positioning the suit pieces?; (iv) How did you feel immersed in the virtual environment?; (v) Do you feel any simulator sickness symptom, like vertigo, nausea, dizziness, difficulty focusing?.

2.1 Experiment 1

36 volunteer healthy subjects, aged between 15 and 45 years (average 25.4 $\pm$ 5.5 years), took part to the experimental session, receiving no reward. They all had normal or corrected to normal vision and had to sign an informed consent.

According to a counterbalanced repeated measure experimental design, participants were asked to com-

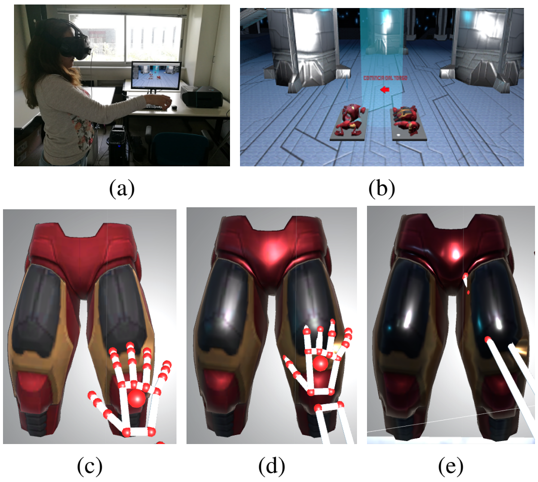


Figure 2: (a) Experiment 2 setup. (b) Game scene. (c) No collision is detected. (d) A collision between hand and suit piece is detected. (e) Hand is closed in a grasping pose.

plete the task with two of the four feedback modalities. Each combination was assigned to 3 subjects and 18 samples for each feedback modality were acquired. A different initial position of shapes is associated to the 4 feedback modality scenes. During each experimental session, first, a demo scene to familiarize with the interface was shown. Then, users performed the task with one of the two feedback modalities. In the main scene, there is a desk with 12 shapes and a block with holes corresponding to the different shapes. Subjects had to grab one object per time and put it in the correct hole. The task ended when all objects were correctly positioned. After a short break, subjects accomplished the task with the second feedback modality.

2.2 Experiment 2

36 volunteer healthy subjects aged between 15 and 50 years (average 25.2 ± 8.1 years) took part to the experimental session, receiving no reward. They all had normal or corrected to normal vision and had to sign an informed consent. Experiment 2 replicates the previous experimental procedure. Each experimental session was composed of an initial scene, where the player read the instructions of the game, and the proper level. When the main scene starts, the player is located between two tables. On the tables there are the suit pieces and in front of him there is a cylinder, where he is instructed to assembly the suit starting from the torso and carrying one piece at a time.

2.3 Experiment 3

90 volunteer healthy subjects played our game. They were between 7 and 53 years old and were divided in three groups based on their age: children (up to 10 years), teenagers (from 11 to 21 years) and adults (from 22 years on). In total 30 children (average 9 ± 1.4 years), 41 teenagers (average 14.4 ± 2.8 years) and 19 adults (average 36.2 ± 12.3 years) took part to our experiment. All of them had normal or corrected to normal vision and signed an informed consent. Permissions were given by parents, when required. Experimental procedure was simplified: all participants accomplished the assembly task, controllers were used for interaction and no feedback were provided.

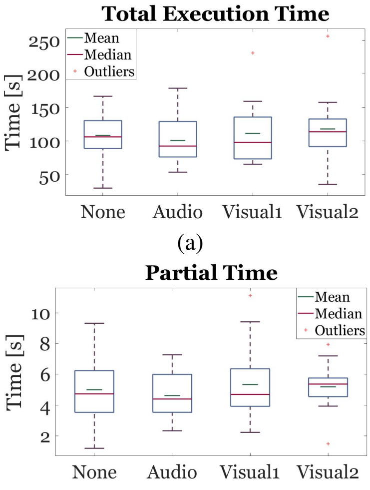
3 RESULTS

In Experiment 1 and 2, quantitative data collected during the experimental session have been analyzed on three levels: first, we compared results obtained with the different feedback modalities; then, we performed a cross comparison between order of execution and feedback modality; finally, we analyzed performances over preference expressed by the users. In each analysis the statistical significance of the differences between samples was estimate through repeated paired sample t-tests.

3.1 Experiment 1

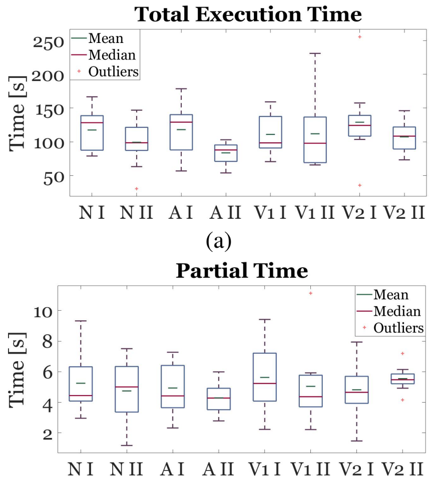
In general, people liked playing our game with all the different feedback modalities and could accomplish the task easily. Preferences were rated as follow: 38% for the *Audio* feedback, asserting that it required less attention and was less distractive, 22% for the *None* modality, 19 % for *Visual 1*, 13 % for *Visual 2* and finally 8 % had no preference. This result confirms our idea that using a sensorial input channel different from the visual one could be a more comfortable way to convey information.

Total completion times referred to feedback modalities (Fig. 3) are comparable (ranging from 100.7 ± 33.5 s to 118.1 ± 44.6 s). However, *Audio* modality is slightly better than the other, while *Visual 2* is the worst. Partial times confirm this results: the one referred to *Audio* are in average the lowest (4.6 ± 1.4 s) and the one referred to *Visual 2* are the highest (5.3 ± 2.3 s). As no statistically significant difference between feedback modalities for total completion time and partial time was found, we can assert



(b)

Figure 3: Experiment 1: comparison of the performances with different feedback. (a) Total completion time (averaged across subjects) and (b) partial positioning time (averaged across subjects and objects).



(b)

Figure 4: Experiment 1: cross-comparison of the performances by considering trial order and different feedback. (a) Total completion time based on feedback and order of execution. (b) Partial completion time based on feedback and order of execution. N = None, A = Audio, V1 = Visual1, V2 = Visual2, while I = first trial and II = second trial.

that feedback do not provide any substantial improvement to performances. The cross comparison analysis highlights a decrease of total completion time, due to learning, but no differences for partial times (see Fig. 4). Statistical analysis ran considering as the dependent variable first the order of execution as the dependent variable then the kind of feedback, found no statistically significant differences, except for total completion time of the first and second trial with *Audio* ($p < 0.05$) and both total and partial time of the second trial with *Audio* and *Visual 2* ($p < 0.05$). Moreover, from Fig. 5, it can be stated that there is no correlation between preferences and performances. T-test run on these data, considering the preference as the dependent variable, highlights no statistically significant difference. However, when designing an interface, it is strongly recommended not ignoring testers preferences. Finally, we analysed the UEQ's 6 scales of evaluation: values between -0.8 and +0.8 represent a neutral evaluation; rates above +0.8 correspond to a positive evaluation, while rates inferior to -0.8 a negative evaluation. In our case, as shown in Fig. 6, all rates have positive values, widely above 0.8, exception made for the *Efficiency* in *Visual 2*, which is 0.842 ± 0.499. In General, *Efficiency* received the lowest ratings, probably because of the problems certain subjects had with hand and gesture recognition, while *Perspicuity* was well rated. People had a moderate to good control over the interface and felt stimulated. However, there is not a clear difference for answers given to the different feedback modalities.

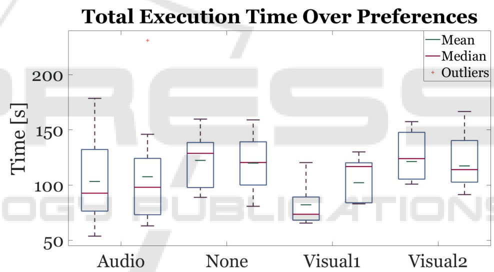


Figure 5: Experiment 1: total completion times over preferences. In each pair, the first is referred to total execution time of people who preferred a certain feedback in the trial with that feedback; the second box is the total time of people who preferred that feedback in the non preferred feedback modality trial.

3.2 Experiment 2

Everybody completed the task successfully and only 9% of participants said they did not appreciate the experience. Half of the volunteers preferred the visual feedback (31% *Visual 1* and 19% *Visual 2*), 14% said the *No feedback* modality was the best one, 11% chose the *Audio* one and 25% had no preference. Although, this result may contradict the one obtained in the previous experiment, in this case, in the scene there were multiple audio sources, when suit pieces collide with the floor or when they are correctly positioned, so multiple sounds conveying different information could be confusing.

Answers to the overview questionnaire (on a Lik-

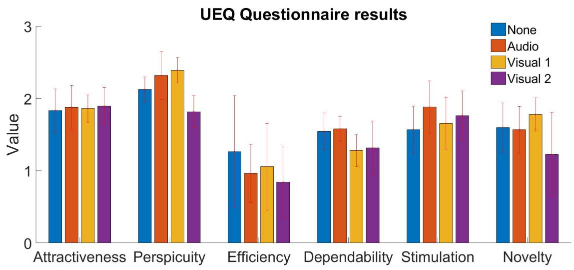
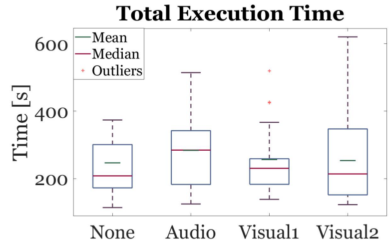
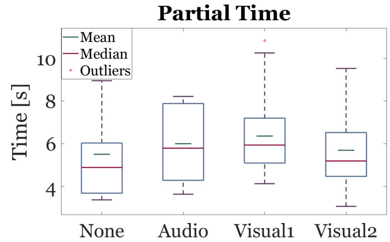


Figure 6: Experiment 1: UEQ results referred to the 4 feedback modalities.

ert scale from 1 to 5) suggest that people found the interaction intuitive (4.2 ± 0.7 points) and the experience immersive (4.3 ± 0.6 points). Nobody felt symptoms of simulator sickness (1.1 ± 0.4 points) and in general participants had more problems on positioning suit pieces (2.9 ± 0.9 points) than on understanding what they corresponded to (1.9 ± 1.1 points). Total execution times referred to different feedback modalities are comparable (see Fig. 7a), and varies in a small range from 246.5 s, in the case of *None*, to 283.7 s, in the case of *Audio*. As shown in Fig. 7b, partial times confirm this trend. Coherently with overview questionnaire answers, people had some difficulties in using the Leap Motion for interaction. One of the major problem is the movement people had to perform in order to carry objects from the table to the cylinder. Due to the limited field of view of the Leap Motion, during head rotations, hands stopped being tracked, thus disappearing from the virtual world, and suit pieces fell on the floor. The t-test never rejected the null hypothesis, i.e. differences among completion times with different feedback are not statistically significant. Considering both the order of execution and the feedback modality, we can notice a learning process more evident in the case of the total completion time in Fig. 8a, as participants have learnt how to build the suit but still have problems with the interaction. Partial times slightly improve in the second trial (see Fig. 8b), except for *Visual* feedback modality. This effect can be due to the fact that, while task is the same, feedback provided varies and participants have to adapt. T-test found that all these differences are not statistically significant. Furthermore, performance are not dependent from preferences (Fig. 9) and differences are not statistically significant. Finally, considering the UEQ (Fig. 10), all the values referred to the different scales are above +0.8. Rates on *Efficiency* and *Dependability*, are the lowest, while *Attractiveness* and *Perspicuity* are the highest. *Visual 1*, in contrast with preferences, received the worst rates, while *Visual 2* the best. *Audio* and *None* are in general comparable.

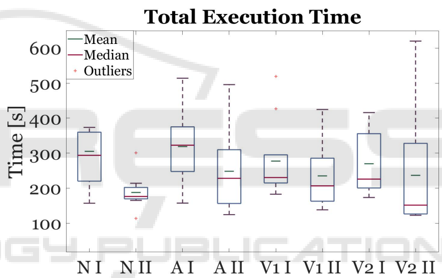


(a)

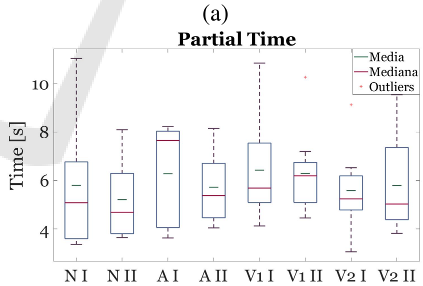


(b)

Figure 7: Experiment 2: comparison of the performance organized on feedback. (a) Total completion time with the various feedback. (b) Partial positioning time with the various feedback.



(a)



(b)

Figure 8: Experiment 2: comparison of the performances organized on feedback and order of execution. (a) Total completion time with the various feedback. (b) Partial positioning time with the various feedback.

3.3 Experiment 3

Quantitative data collected during the experimental acquisition have been divided in three groups based on player's age (Children, Teenagers and Adults). Performances, in fact, differs between groups. Considering the total execution time, shown in Fig. 11a,

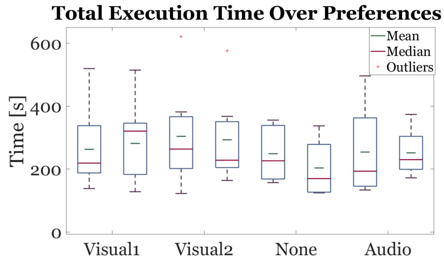


Figure 9: Experiment 2: total completion times over preferences. In each pair, the first is referred to total execution time of people who preferred a certain feedback in the trial with that feedback; the second box is the total time of people who preferred that feedback in the non preferred feedback modality trial.

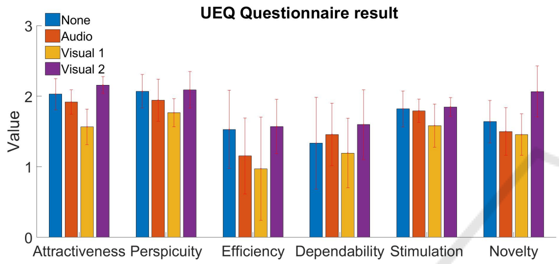


Figure 10: Experiment 2: UEQ results referred to the 4 feedback modalities.

we can notice that children had more difficulty in completing the task (390.9 ± 180.9 s) with respect to teenagers (225.1 ± 129.8 s) and adults (250.4 ± 135.3 s), whose performances are almost comparable. T-test analysis highlights a statistically significant difference between Children and Teenagers ($p < 0.001$) and between Children and Adults ($p < 0.01$), while Teenagers and Adults are not statistically significant different.
Partial times, instead, are very similar, ranging from 3.1 s for Children and 3.3 s for Adults, and differences are not statistically significant (see Fig. 11b). This means that performances are more influenced by the players' cognitive skills and strategy adopted than by the intuitiveness or efficiency of the interaction: children tended to have more difficulties in recognizing the single pieces or in understanding instructions, as confirmed by the overview questionnaire. For participants it was more difficult understanding what pieces corresponded to (3.9 ± 1.1 points) than actually positioning them (2.9 ± 1.2 points). Moreover, in general, everyone liked the game and had no simulator sickness (1.3 ± 0.9 points), commands have been considered intuitive (4.2 ± 1.1 points) and environment immersive (4.7 ± 0.8 points). Finally, comparing the two complex task with the Leap Motion and the controller, we can see that total times are comparable in the two cases, while partial times are slightly better in

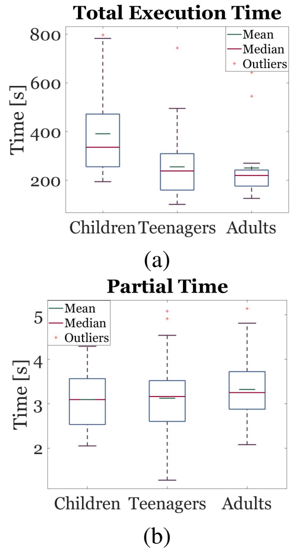


Figure 11: Experiment 3: (a) total completion times and (b) partial times results divided per participant age.

Experiment 3, confirming the hypothesis that worsen results in Experiment 2 are affected more by the interaction device used, than by the cognitive load of the task.

4 CONCLUSIONS

In this paper, we have described three different experiments aiming to investigate the interaction in immersive virtual environments, considering different input devices, feedback provided and task complexity. In particular, as the absence of haptic feedback has been proven to be an impairment when interacting with virtual objects (Mine et al., 1997), we wondered whether other sensory feedback, e.g. visual and audio, could substitute it and contribute to improve performances.

Thus, in Experiment 1, once ascertained Leap Motion usability for VR applications (Gusai et al., 2017), we decided to add feedback and compared an audio feedback, two visual feedback and a control condition. Total completion and partial times in the four cases were similar. However, participants expressed a preference for *Audio* feedback, maybe due to the fact that it involves a sensory channel which is not directly addressed in VR.
In Experiment 2, we decided to introduce a more complex task in a more structured environment. In fact, even if Leap Motion has been proven to be an adequate solution for simple interactions, its usability for complex tasks is still under debate. Even in this case, feedback seems not to affect performances and there is not a clear preference among users. Probably

because of the soundtrack and the collision sounds, already present in the scene, *Audio* did not play a fundamental role.
Finally, we replicated the complex task using controllers for interaction, in order to understand if results in Experiment 2 were related to the task cognitive load or to difficulties in managing the interaction. Results confirms an influence of the selected interface over performances .

In conclusion, if we consider the research questions (see Section 1.2), we can conclude as follows.

- Considering that performances with or without feedback in Experiment 1 and 2 were similar and not significantly different, we can state that in 3D immersive virtual environment feedback can not significantly improve performances, contrary to results obtained for non immersive virtual environment (Chessa et al., 2016).
- Although results show no significant improvement in performances, they do not exclude the beneficial effect of haptic feedback.
- Notwithstanding, there is no correlation between performances and preferences, as found in literature previous studies.
- While task complexity seems playing a fundamental role on total execution time, there is no effect on partial times, thus it has a role on participants cognitive load, but not on the interaction to be used.

In future, it would be interesting to design an interface based on the combined use of haptic feedback and Leap Motion and test it in both cases, in order to assess its usability and understand till what extent performances with touchless interfaces can be improved in order to be as effective as controllers based one.

REFERENCES

Argelaguet, F. and Andujar, C. (2013). A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136.

Augstein, M., Neumayr, T., and Burger, T. (2017). The role of haptics in user input for people with motor and cognitive impairments. *Studies in health technology and informatics*, 242:183–194.

Bassano, C., Solari, F., and Chessa, M. (2018). Studying natural human-computer interaction in immersive virtual reality: A comparison between actions in the peripersonal and in the near-action space. In *VISIGRAPP (2: HUCAPP)*, pages 108–115.

Bowman, D. A., McMahan, R. P., and Ragan, E. D. (2012). Questioning naturalism in 3D user interfaces. *Communications of the ACM*, 55(9):78–88.

Chessa, M., Matafu, G., Susini, S., and Solari, F. (2016). An experimental setup for natural interaction in a collaborative virtual environment. In *13th European Conference on Visual Media Production (CVMP13)*, pages 1–2.

Cuervo, E. (2017). Beyond reality: Head-mounted displays for mobile systems researchers. *GetMobile: Mobile Computing and Communications*, 21(2):9–15.

Frohlich, B., Hochstrate, J., Kulik, A., and Huckauf, A. (2006). On 3D input devices. *IEEE computer graphics and applications*, 26(2):15–19.

George, R. and Blake, J. (2010). Objects, containers, gestures, and manipulations: Universal foundational metaphors of natural user interfaces. *CHI 2010*, pages 10–15.

Gusai, E., Bassano, C., Solari, F., and Chessa, M. (2017). Interaction in an immersive collaborative virtual reality environment: A comparison between leap motion and htc controllers. In *International Conference on Image Analysis and Processing*, pages 290–300. Springer.

Laugwitz, B., Held, T., and Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*, pages 63–76.

McMahan, R. P. (2011). *Exploring the effects of higher-fidelity display and interaction for virtual reality games*. PhD thesis, Virginia Tech.

Mine, M. R., Brooks Jr, F. P., and Sequin, C. H. (1997). Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 19–26. ACM Press/Addison-Wesley Publishing Co.

Schrepp, M., Hinderks, A., and Thomaschewski, J. (2014). Applying the user experience questionnaire (UEQ) in different evaluation scenarios. In *International Conference of Design, User Experience, and Usability*, pages 383–392.

Zhang, F., Chu, S., Pan, R., Ji, N., and Xi, L. (2017). Double hand-gesture interaction for walk-through in VR environment. In *Computer and Information Science (ICIS), 2017 IEEE/ACIS 16th International Conference on*, pages 539–544. IEEE.