3D Object Arrangement for Novice Users: the Effectiveness of Combining a First-Person and a Map View

Luca Chittaro Roberto Ranon Lucio Ieronutti

Human-Computer Interaction Lab
Department of Mathematics and Computer Science
University of Udine
Via delle Scienze, 206
33100 Udine, Italy

firstname.lastname@dimi.uniud.it

ABSTRACT

Arranging 3D objects in Virtual Environments can be a complex, error prone and time consuming task, especially for users who are not familiar with interfaces for 3D navigation and object manipulation. In this paper, we analyze and compare novice users' performance on 3D object arrangement tasks using three interfaces that differ in the views of the 3D environment they provide: the first one is based only on a first-person view; the second one combines the first-person view and a map view in which the zoom level is manually controlled by the user; the third one extends the second with automated assistance in controlling the map zoom level during object manipulation. Our study shows that users without prior experience in 3D object arrangement prefer and actually benefit from having a map view in addition to a first person view in object arrangement tasks.

Keywords

3D manipulation, virtual environments, user study, experimental evaluation.

1. INTRODUCTION

In many popular applications for building Virtual Environments (VEs), such as game level building, creation of 3D virtual exhibitions, and virtual home/interior design, *object arrangement* is one of the major tasks. Object arrangement involves: (i) 3D navigation (finding and reaching the place where a given object has to be positioned), (ii) object selection, and (iii) object manipulation (properly positioning, orienting and scaling the object) [1]. Even in applications that are not meant for 3D modelers or CAD engineers, arranging 3D objects inside a VE can still be a complex, error prone and time consuming task for users with minimal or no experience in 3D object manipulation (hereinafter, we refer to this category as *novice users*).

Several solutions for supporting navigation or manipulation tasks are based on providing the user with multiple views of the VE. However, interacting with (and relating) multiple views can require a considerable cognitive effort, as anyone who has tried to use the typical 4-views layout adopted by 3D modeling software can testify. For this reason, approaches meant for children or novice users (which we discuss in Section 2) propose a single (first-person or third person) view of the VE. While this solution seems to make sense for novice users, it might make arranging objects more difficult, because users can see only a limited part of the VE.

For navigation purposes, an electronic map of the VE is a common addition to the single view approach. Compared to other multi-view solutions (e.g., World in Miniature [2]), a map can provide *global* (showing the entire VE) as well as *local* (showing parts of the VE in more detail) information with minimal interaction effort (i.e., setting the map zoom level). However, while different studies (e.g., [3][4]) demonstrate the effectiveness of using maps for supporting navigation in VEs, or compare maps to 3D views for object positioning tasks [5], no studies have been specifically targeted at evaluating the benefits for novice users of introducing maps for object arrangement purposes. The primary goal of this paper is to investigate such issue, while our secondary goal is to evaluate two alternative solutions for controlling the zoom level in the map view. In particular, we analyze and compare novice users' performance on 3D object arrangement tasks in three conditions: the first one provides only a first-person view; the second one combines the first-person view and a map view in which the zoom level is manually controlled by the user; the third one provides the same two views of the second one but adds automated assistance in controlling the map zoom level during object manipulation. While it would be also interesting to perform the same study with more experienced users, in this paper we concentrate on novice users since we are interested in investigating the main barriers that prevent those users from autonomously building VEs and creating user-generated 3D contents.

The paper is structured as follows. Section 2 surveys previous research and applications that aim at simplifying the construction of VEs for novice users. In Section 3, we describe in detail the interfaces considered in the comparative analysis.

Section 4 presents the user study we carried out. Section 5 and 6 report and discuss the obtained results.

2. RELATED WORK

The problem of how many and which views are more effective for facilitating interaction with VEs has been dealt with by several authors. Some of them specifically proposed applications for VE creation by novice users. Wang and colleagues [6] developed the *Kids Movie Creator* (KMC), a tool that allows children (aged 7 to 12) to create interactive VEs. KMC uses a single, third-person view; the user controls the avatar position using a walk navigation mode through arrow keys, while mouse drags allow her to see the avatar from different perspectives or control its orientation. *LEGO Digital Designer* (LDD) [7] is a tool to create VEs by positioning and combining elementary 3D models representing LEGO bricks. The VE is displayed with a single view, using an examine navigation mode controlled through mouse drags, while the mouse wheel is used to control the zoom level. *Alice* [8] is a 3D programming environment designed for undergraduates with no 3D graphics and programming experience that allows to build VEs and program their behavior. In Alice, navigation is based on a single-view, fly mode. Bowman et al. [9] instead propose the use of a first-person view together with a map view in *Virtual Habitat*, an educational application that allows users to build virtual zoo exhibits.

Other research focuses on interface techniques, not specifically devoted to a class of applications or users. Researchers have experimented with 2D [10] and 3D [2][11] map-based solutions, or with secondary views that provide an additional viewpoint to display the surroundings of the user [12]. While most of these approaches focus on supporting navigation, some of them [2][12] can also be used for object manipulation. There are also proposals that rely on an arbitrary number of views [13] which can be created and deleted by the user at her will.

Some studies compare the effectiveness of using different view configurations. For example, Tory et al. [14] compare different combinations of 2D and 3D views for object orientation and relative positioning estimation, concluding that the combination of 2D and 3D displays has better performance than 2D or 3D alone. In [5], a 3D perspective view and a 2D map view are compared in object positioning tasks, with mixed results.

Finally, it is also worth noting that using multiple views to improve interaction is an issue that arises and is of interest also to other fields, for example the work on coordinated multiple views in the Information Visualization community (e.g., [15]), and the work on multiple frames of reference for collaborative object arrangement in the CSCW community (e.g., [16]).

3. THE CONSIDERED INTERFACES

In our study, we compare the effectiveness of using three different interfaces for object arrangement, which we will call First-person view (*FPV*), First-person plus Map view (*FPV+M*), and FPV+M with Assisted Zoom Control (*FPV+MAZ*). We made this choice because we (i) focus on novice users of Desktop VEs and this rules out the solutions in the literature that are based on immersive displays or special input devices, (ii) the use of single 3D views or 3D views plus a map has been the solution of choice for novice users and the considered task in the literature.

In the following, we describe in detail the three considered interfaces. In FPV, the user sees a single, first-person view of the VE, while in FPV+M the screen space is divided in two parts, as in Figure 1: the left part (65% of the screen space) displays the first-person view, while the right part displays a map of the VE. The map is:

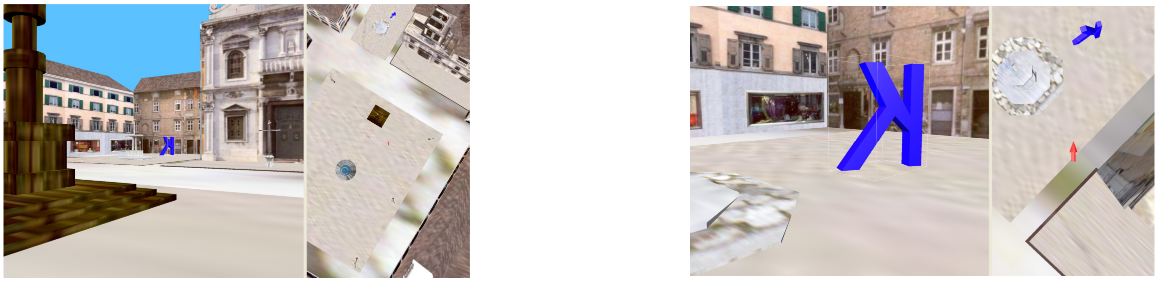
- a *you-are-here map*, in which an oriented arrow is used for representing current user position and orientation in the VE;
- a *forward-up map*, i.e., the orientation of the map dynamically changes such that the upper part of the map shows what is in front of the user in the first-person view; for the kind of search tasks we consider in our experiment, it has been shown that a forward-up map is better or not significantly different than a north-up map [17];
- *centered on the user*, i.e., when the user moves through the VE, the map moves as well to maintain the arrow at the center of the view;
- *zoomable*, allowing users to obtain a global map of the VE (by zooming out) as well as a more detailed, local map of the places and objects near her current position (by zooming in).

Moreover, in the map view, when the user is inside a structure (e.g., a building), a cutting plane is inserted just above the user's head and parallel to the floor where the user stands. In this way, any object that may stand above the user is not drawn and does not obstruct the view of the user's surroundings.

While in FPV+M the user has to manually control the zoom level of the map, in FPV+MAZ the application dynamically sets the maximum zoom level of the map such that, as soon as an object is selected in the first-person view, the selected object is entirely included in the map. Figure 1 depicts the effects of the assisted zoom control in the case of an object that is respectively far from (Figure 1a) and near (Figure 1b) the current user position. With this solution, we aim at reducing the effort required for manually controlling the zoom level during object manipulation. Additionally, since such assisted control ensures that a suitable level of zoom is adopted during object manipulation, the user could be able to effectively use the map view both for coarse (if the object is far from the user position) and fine (if the object is near the user) manipulations. In FPV+MAZ, the user can still manually control the zoom level when no object is selected.

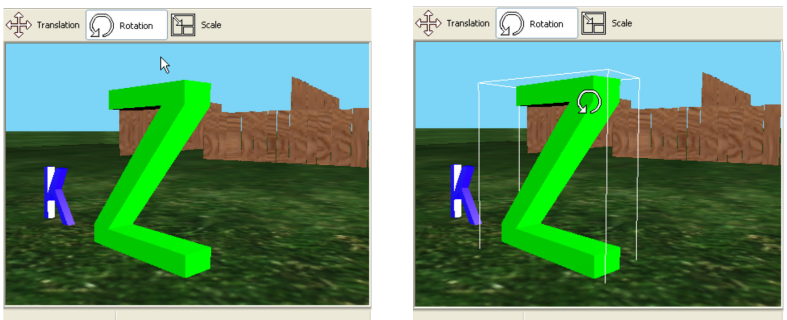
In our study, we consider manipulation tasks involving 3DOF for translations, and 1DOF for rotations and scale operations. In particular, users can rotate each object along its vertical axis, and scale it preserving its original proportions.

The reduction of the number of DOF for rotation and scale operations is a reasonable simplification in the context of applications designed for novice users. For example, only 1DOF is available to the user for scaling objects in KMC [6], and 1DOF is available for rotations in Alice [8].



(a) (b)

Figure 1. Views in FPV+MAZ when a 3D object shaped as a letter "K" is selected: (a) the object is far from the current user position (b) the object is near the current user position.



(a) (b)

Figure 2. (a) The mouse pointer is not positioned over an object, (b) when the mouse pointer is positioned over the object, the bounding box of the object and the shape of the mouse pointer indicate the possibility of manipulation.

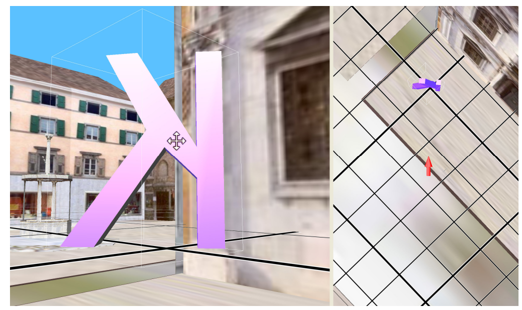


Figure 3. During translation of an object, a grid parallel to the xz-plane and aligned with the bottom of the object is displayed.

From the point of view of navigation and manipulation controls, we have adopted commonly used solutions, as we describe in the following subsections.

3.1 Navigation Control

In the three experimental conditions, users move through the VE in a walk navigation mode, where collision detection, gravity and terrain following are enabled. Navigation is based on traditional (game-like) controls; keyboard and mouse are used to control respectively the position and orientation of the user into the VE. In particular, the up, down, left and right arrow keys are used for moving user position respectively forward, backward (in the user's current viewing direction) and sideways (without changing user's orientation).

Mouse drags (with any mouse button) are used for controlling the avatar orientation when they do not start over a manipulable object. In particular, the horizontal and vertical components of the mouse movement are used to control respectively the yaw and pitch of the avatar head. As soon as the avatar changes its horizontal orientation, the forward-up map in FPV+M and FPV+MAZ is rotated accordingly.

We did not allow the possibility of dragging the avatar in the map view (as in World-In-Miniature [2] and similar systems) because locomotion speeds in FPV+M and FPV+MAZ would be then much higher than those in FPV and task completion times would be affected by an uncontrolled factor.

3.2 Manipulation Control

In all three conditions, the type of manipulation (translation, rotation or scale) is chosen by clicking on the corresponding icon available in the upper part of the first person view window (see Figure 2(a) and Figure 2 (b)).

To apply the currently chosen manipulation to an object, the user has to start a drag action over the object. Two visual aids highlight the possibility of interacting with an object when the mouse pointer is over it: (i) the pointer changes its shape according to the currently selected manipulation mode, and (ii) the bounding box of the object is displayed. For example, in Figure 2(a), rotation is the currently selected manipulation mode but the mouse pointer is not positioned over an object. In this case, mouse drags control user orientation. In Figure 2(b), rotation is the currently selected manipulation mode and the mouse pointer is over a manipulable object. Table 1 describes in detail the controls for selecting, translating, rotating and scaling objects, as well as for zoom level manipulation.

To make object positioning easier, a grid parallel to the xz-plane and aligned with the bottom of the bounding box of the manipulated object is displayed during the translation (see Figure 3). As the vertical position of the currently manipulated object changes, the position of the grid is updated accordingly. The grid is aimed at providing additional visual cues to understand the correct position of the selected object in 3D space, especially during vertical positioning.

Table 1. Controls for selecting. translating, rotating and scaling objects, as well as for zoom level manipulation.

Mode	View	Action	Effect
Selection	First-person or Map	Mouse Click (left or right button)	Selects the object pointed by the mouse cursor. However, any mouse drag implicitly selects an object and starts a manipulation with it.
Translation	First-person	Mouse drag (left button)	Moves the object along the world *xz*-plane without changing its vertical position. The vertical component of the mouse drag moves the object closer/away from the user, the horizontal component moves the object left/right.
Translation	First-person	Mouse drag (right button)	Moves the object along the plane which passes through the centre of the object and orthogonal to the current viewing direction of the user.
Translation	Map	Mouse drag (left button)	Moves the object along the world *xz*-plane without changing its vertical position.
Translation	Map	Mouse drag (right button)	Moves the object along the world *y*-axis. The vertical component of the mouse drag is mapped into a vertical translation of the object.
Rotation	First-person or Map	Mouse drag (left or right button)	Rotates the object around the world *y*-axis. The horizontal component of the mouse drag is mapped into a rotation around the world *y*-axis.
Scale	First-person or Map	Mouse drag (left or right button)	Uniformly scales the object. The vertical component of the mouse drag modifies the size of the object.
Zoom	Map	Mouse wheel scroll	Increases or decreases the zoom level of the map. Scroll up and down actions correspond respectively to zoom in and zoom out operations.

4. EXPERIMENTAL EVALUATION

Our hypotheses in the present study are the following:

- the availability of the map view should reduce the time and the number of actions required to arrange objects in a VE.
- the assisted zoom functionality should alleviate part of the effort of manually controlling the zoom level during manipulation tasks, and therefore result in less time to complete the task and less number of zoom operations required.

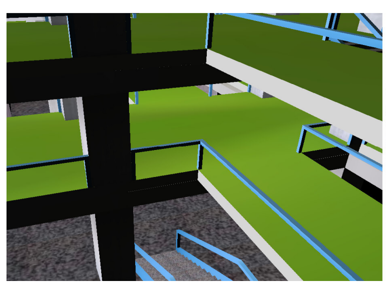
4.1 Participants and Task

The evaluation involved a sample of 16 subjects, 7 male and 9 female. Subjects were volunteers recruited by email or direct contact. Age ranged from 20 to 49, averaging at 34. The majority of subjects (75%) had an education in arts and humanities; the others had a scientific background. They had different levels of familiarity with computers: 3 subjects use a computer less than 5 hours per week, 2 from 6 to 15 hours, 6 from 16 to 30 hours, and the remaining 5 subjects use the computer more than 30 hours per week; 12 subjects had never played 3D videogames, while the remaining 4 subjects play them less than one time per month. All subjects had no experience with any kind of game level editor, 3D modeling tool or CAD system.

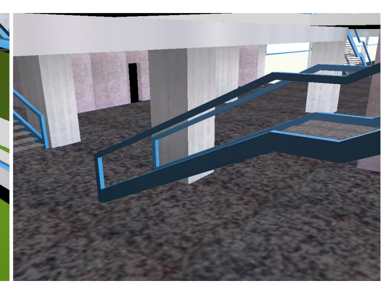
Subjects experimented the three interfaces in two VEs with very different structure. The first one (OUTDOOR) is an outdoor and open environment (a city square). This VE includes highly distinctive landmarks to aid orientation, e.g., the central fountain or the church (see Figure 4a) and has no branching points. The second VE (INDOOR) is a multi-floor, indoor environment (see Figure 4b and 4c), with no particularly distinctive landmarks, and a considerable number of columns and walls that reduce, with respect to the previous VE, freedom of movement and visual access. Moreover, the INDOOR environment has several branching points. As a result, navigation becomes more difficult and occlusion problems occur more frequently. The rationale for studying two VEs is that we want to cover typical exterior and interior settings.

For each environment and interface condition, participants performed a task composed by two manipulation sub-tasks, one involving a “K”-shaped object, the other a “Z”-shaped object. The object to be manipulated was initially positioned in front of the user, initially not selected, and visible in all available views. The user had to manipulate the object to reach a target configuration indicated by a semitransparent copy (ghost) of the object placed in the target position, orientation and scale (see Figure 5). In OUTDOOR, the ghost was visible in the initial view, but quite far away from the user initial position (roughly, the user started on one side of the VE, and the ghost was on the other side). In INDOOR, the ghost was not visible from the initial view, and it was positioned on a different floor with respect to the user’s initial position. However, the user was informed about the position of the ghost by means of printed paper screenshots showing also her initial position through

(a)



(b)



(c)

Figure 4. (a): a part of the OUTDOOR VE; (b) and (c): parts of the multi-floor, INDOOR VE.

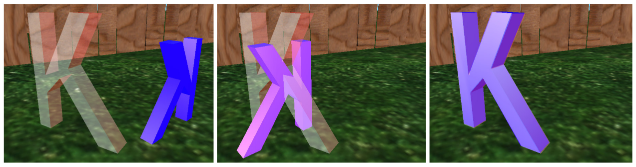

(a) (b) (c)

Figure 5. Final steps of object positioning: (a) the object and the ghost are clearly separated, (b) the object is translated over the ghost, and (c) the object is scaled, rotated and translated to match the ghost.

transparencies and cutaways.

Each single sub-task ended when the subject felt she had completed it or believed that she was not able to further improve the match between object and ghost.

4.2 Procedure

Following a within-subjects design, every subject was tested in every experimental condition. As a result, there were 6 tasks for each subject, 3 for each of the two VEs. Subjects were initially asked to fill a questionnaire to collect demographic information and data about experience in computer use, 3D games and modeling tools. Then, subjects were orally instructed about the task to be performed.

In each experimental condition, subjects were initially allowed to spend unlimited time in a training environment until they felt familiar with navigation and manipulation controls, application toolbar (for the selection of the manipulation mode) and shapes of objects to be manipulated. The environment used for this training phase was characterized by a combination of the salient features of the two VEs used for the experiment. In particular, it represented a small multi-floor building with an external garden. During training, two manipulation tasks were given to the subject, each one meant to familiarize the subject with different VE characteristics.

After completing the training phase, users carried out the experimental task in the two VEs and were primed by simply telling them to complete the task as accurately as possible. To make it easier to recognize the objects to be manipulated and identify the locations of ghosts, the subjects were provided with a color printed sheet, containing images of such objects into the considered VE.

To avoid learning effects due to repetitive testing, the order of conditions and the order of VE were counterbalanced. Moreover, each subject did not perform the same exact manipulation task twice. To this purpose, we defined 3 equivalent but slightly different tasks for each VE, i.e. the initial and target positions of the objects varied. However, all tasks had the same complexity in terms of: (i) euclidean, angular and scale distances between the initial and target configurations; (ii) visibility of the target configuration from the user initial position; (iii) number of floors between the initial and target configurations in INDOOR. The order of the 3 tasks changed independently of the order of conditions.

At the end of the experiment, subjects were asked to fill a questionnaire to collect subjective preferences and comments. In particular, users had to rate the three interfaces from the best to the worst, specifying also main strengths and weaknesses of each.

4.3 Experiment variables

In our experiment, the independent variable was the type of view (FPV, FPV+M, FPV+MAZ), while the dependent variables were:

- task completion time, i.e., time required to complete the task, defined as the time elapsed between the first and last user interaction (navigation or manipulation);
- manipulation time, i.e., time spent in manipulating the objects to complete the task;
- navigation time, i.e., the time spent in navigating the VE;
- number of manipulations, i.e., number of translations, rotations and scale operations required for completing the task;
- manipulation accuracy, i.e., distance between the target and final object configurations, defined as a combination of translation (measured in meters), rotation (measured in degrees) and scale (measured in percentage) distances.

For measuring dependent variables, during each task we automatically recorded the following data:

- manipulations: the start time, end time, duration and type (translation, rotation or scale) of each mouse drag operation to manipulate an object.
- object configuration: the position, orientation and scale of each object to be manipulated, recorded every 250 milliseconds;
- user navigation: the position and orientation of the user in the VE recorded every 250 milliseconds.

5. RESULTS

All the data analyses we report in this section for task completion times, number of manipulations, and manipulation accuracy, were preceded by a Kolmogorov-Smirnov test of normality. In those cases where the test did not reveal deviations from the normal distribution, we proceeded with a one-way analysis of variance (ANOVA). In a few cases, the Kolmogorov-Smirnov test revealed a moderate degree of non-normality due to the presence of outliers. Since these outliers were legitimate values and were not the result of some kind of mistake or mishap, in those cases we performed a logarithmic transformation of the data to reduce their impact and make the distribution more symmetric before carrying out the ANOVA. In those cases where the ANOVA returned a $p<.05$, we then used Tukey's post-hoc test to directly compare pairs of conditions.

5.1 Task completion times

For the time spent by subjects (Figure 6a) to complete the task in the two VEs, the ANOVA pointed out a significant effect ($F(2, 30) = 5.835$, $p < 0.01$) and the direct comparisons between pairs of conditions highlighted that FPV+M required significantly less time than FPV ($q = 4.425$, $p < 0.05$) and FPV+MAZ ($q = 3.892$, $p < 0.05$).

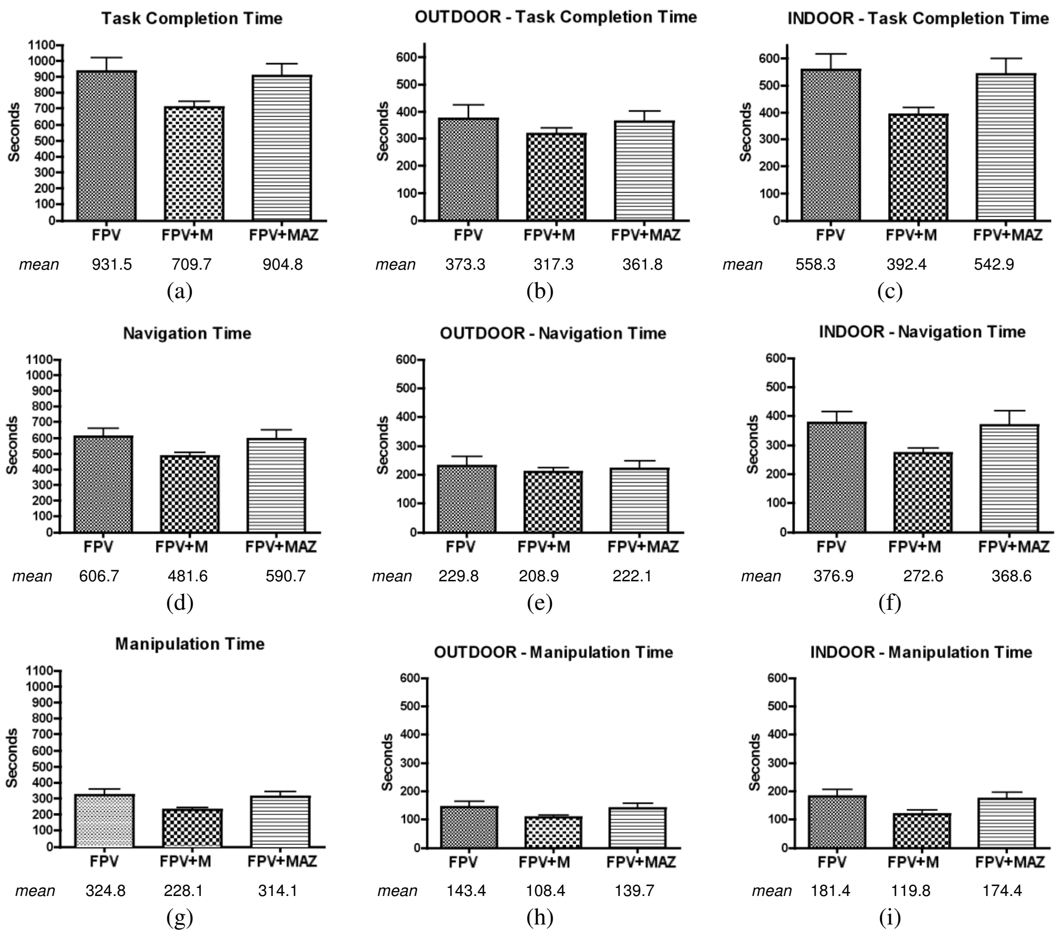


Figure 6. Mean task completion, navigation and manipulation times.

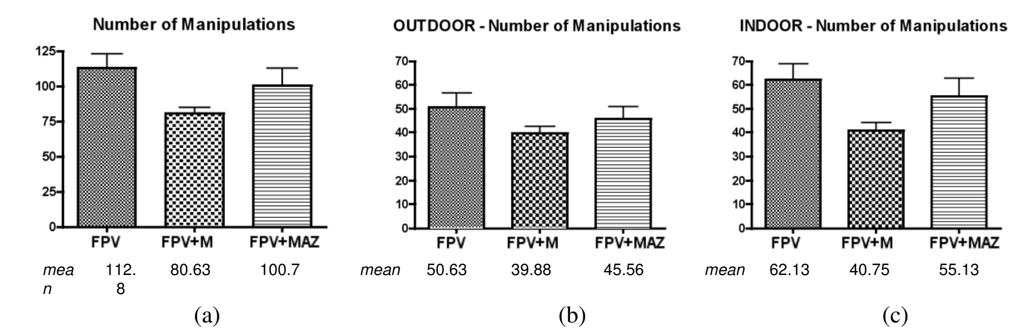


Figure 7. Mean number of manipulations.

By separately analyzing the time required for completing the task in OUTDOOR (Figure 6b) and INDOOR (Figure 6c), we learned that carrying out the task in INDOOR contributed the most to the difference in times. The ANOVA highlighted no significant effect in the case of OUTDOOR, and a significant effect ($F(2,30) = 5.365$, $p < 0.05$) in the case of INDOOR, for which Tukey's test showed that FPV+M required significantly less time than FPV ($q = 4.176$, $p < 0.05$) and FPV+MAZ ($q=3.825$, $p < 0.05$).

Then, we separately analyzed the time spent by subjects in navigating the VEs (Figure 6d) and the time spent for object manipulation (Figure 6g). For navigation times, the ANOVA highlighted a significant effect of the interface ($F(2, 30) = 4.885$, $p < 0.05$), and Tukey's test pointed out that navigation time in FPV+M was significantly less than FPV ($q = 4.061$, $p < 0.05$) and FPV+MAZ ($q = 3.542$, $p < 0.05$). The analysis of navigation logs allowed us also to discover that differences in navigation time among subjects mainly concerned navigation speed rather than the length of the path followed to complete tasks. Also in this case, it was INDOOR that contributed the most to the result: the separate analysis of navigation times for INDOOR (Figure 6e) and OUTDOOR (Figure 6f) pointed out no significant effect for OUTDOOR and a significant effect for INDOOR

($F(2,30)=3.480$, $p < 0.05$). Post-hoc analysis pointed out that FPV+M required significantly less time in navigating INDOOR than FPV ($q = 3.617$, $p < 0.05$).

We obtained analogous results for the time spent by users in object manipulation (Figure 6g). The ANOVA highlighted a significant effect ($F(2,30) = 4.169$, $p < 0.05$) of the interface, and the direct comparison revealed that FPV+M required significantly less time in object manipulation than FPV ($q=3.594$, $p < 0.05$). Also in this case, it was INDOOR that contributed the most to the result: the separate analysis of INDOOR (Figure 6h) and OUTDOOR (Figure 7i) revealed a significant effect of the interface ($F(2,30) = 3.918$, $p < 0.05$) only for INDOOR, and Tukey's test pointed out that FPV+M required significantly less time in object manipulation than FPV ($q = 3.560$, $p < 0.05$).

5.2 Number of object manipulations

For the number of object manipulations (translation, rotation and scale) required for completing the task (Figure 7a), the ANOVA revealed a significant effect ($F(2,30) = 5.888$, $p <0.01$), and the direct comparison pointed out that FPV+M required a significantly smaller number of manipulations than FPV ($q=4.804$, $p < 0.01$).

We also analyzed separately the number of manipulations in OUTDOOR (Figure 7b) and INDOOR (Figure 7c). The ANOVA highlighted no significant effect for OUTDOOR and a significant effect for INDOOR ($F(2,30) = 5.132$, $p < 0.05$), Tukey's post-hoc test pointed out that FPV+M required a significantly smaller number of manipulations than FPV ($q=4.519$, $p < 0.01$).

To study how accurate the result of manipulations was, we derived from logged data the distances between the target configurations and the configurations obtained by the users with the manipulated objects, in terms of translation (measured in meters), rotation (measured in degrees) and scale (measured in percentage). Tasks carried out with FPV+MAZ generally gave more accurate results than the other two interfaces, but statistical analysis did not point out a significant effect. In particular, ANOVA did not reveal a significant effect for translation ($F(2,30) = 0.95$, $p = 0.39$) or rotation ($F(2, 30) = 0.72$ $p = 0.48$), while for scale operations it highlighted a significant effect ($F(2,30) = 0.20$ $p < 0.05$), but the direct comparison pointed out no significant differences between pairs of conditions.

5.3 Subjective preferences

Users where asked to rate the three interfaces from the best to the worst. For the analysis, we assigned a score of 3, 2 and 1 respectively to the first, second and third ranked interface. Figure 8 shows the means of user preferences for each interface. We analyzed the data with Friedman's test, that revealed a significant effect ($T = 6.125$, $p < 0.05$) and Dunn's test for post-hoc analysis, that showed a statistically significant difference between FPV+MAZ and FPV ($p < 0.05$).

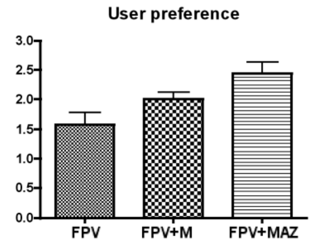


Figure 8. Means of user preferences.

6. DISCUSSION AND CONCLUSION

Overall, most of the statistically significant differences between pairs of conditions concerned FPV and FPV+M. In particular, the analysis pointed out that FPV+M requires significantly less time and number of manipulations than FPV. Moreover, subjective data showed that users preferred to have a map view. Both findings are consistent with the first hypothesis of the study. The evaluation also pointed out that the benefits of using FPV+M were significant in the case of INDOOR, but not significant in the case of OUTDOOR.

By observing the subjects during the experiment, we noted that they used the map mainly for two purposes. First, the map was used during navigation to see parts of the surrounding environment that were not visible in the first-person view (e.g., parts behind the user). Several users reported in the post-experiment questionnaire that this was helpful when, using the left, right and down arrow keys, they moved towards parts of the VE that were out of their field of view. Second, we noticed that the map was often used for object manipulation, especially to perform the first rough translations and rotations.

Although the subjective preferences indicated FPV+MAZ as the best interface, users' actual performance was not consistent with this ranking, and the second hypothesis of our study was thus not confirmed. In the post-experiment questionnaire, most users expressed positive remarks concerning the assisted zoom control, indicating that it gave them the feeling of being more supported during object manipulation, and alleviated the load of manually controlling the zoom level. Indeed, the average number of zoom operations with FPV+MAZ was lower than in FPV+M, although the difference is not statistically significant. However, the total number of manipulations performed and time completion time with FPV+MAZ were significantly higher. We have two hypotheses to explain this. First, in some cases FPV+M could provide a wider view than FPV+MAZ, meaning that users could translate objects farther in the map view. Since the automatic zoom view did not necessarily include the target location, users might have had to translate the object several times in order to get it to the target. The second hypothesis is that, since the assisted zoom provided the most detailed possible view of the object in the map, it encouraged users to reach a better level of accuracy at the expense of more manipulations. As noted above, arrangement accuracy was indeed greater with FPV+MAZ than in the other conditions, although without statistical significance.

The results of our study provide practical indications for the design of interfaces aimed at enabling non-technical users to build VEs. More specifically, our study shows that users without prior experience in 3D object arrangement prefer and actually benefit from having a map view in addition to a first person view in object arrangement tasks. In particular, the FPV+M interface effectively decreased both the task completion time and the number of manipulations required. On the contrary, the automatic zoom feature we proposed with the aim of obtaining even better results turned out to be detrimental to user performance.

One possibility to improve the assisted zoom mechanism could be to use it also during object translation or scaling in the map to ensure that the object is always clearly visible. For example, when the object is dragged outside the portion of the VE currently displayed by the map view, the system could automatically zoom out to keep the object visible. Note that panning on the map would not be a solution because the user position is kept in the center of the map. Another possibility that could be worth evaluating is to reduce the number of DOF available to users. For example, we noted during the experiment that the effort required for aligning an object with the terrain was considerable. Simplifying the interaction by constraining object movements to the surfaces of the VE might thus considerably reduce the time spent for object positioning in some applications (e.g. virtual interior design/home planning).

Acknowledgments

The authors acknowledge the financial support of the Italian Ministry of Education, University and Research (MIUR) within the FIRB project number RBIN04M8S8.

7. REFERENCES

[1] Bowman, D., Kruijff, E., Laviola, J., and Poupyrev, I. 2004. “3D User Interfaces: Theory and Practice”. Addison-Wesley Professional.

[2] Stoakley, R., Conway, M.J., and Pausch, R. 1995. Virtual reality on a WIM: Interactive worlds in miniature. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 1995), ACM Press, New York, 265–272.

[3] Darken, R.P., and Sibert, J.L. 1996. Wayfinding Strategies and Behaviors in Large Virtual Worlds. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 1996), ACM Press, New York, 142-149.

[4] Ruddle, R.A., Payne, S.J., and Jones, D.M. 1999. “The Effects of Maps on Navigation and Search Strategies in Very large-scale Virtual Environments”. Journal of Experimental Psychology: Applied, 5(1), 54-75.

[5] St. John, M., Smallman, H.S., Bank, T.E., and Cowen, M.B. 2001. Tactical Routing Using Two-Dimensional and Three-Dimensional Views of Terrain. Technical Report 1849, SSC San Diego Technical Reports.

[6] Wang, T., Li, X., and Shi, J. 2007. An Avatar-Based Approach to 3D User Interface Design for Children. In Proceeding of IEEE Conference on 3D User Interfaces, IEEE Computer Society Press, Los Alamitos, 155-162.

[7] Lego, 2008. Lego Digital Designer, http://ldd.lego.com.

[8] Conwaym, M., Audia, S., Burnette, T., Cosgrove, D., Christiansen, K., et al. 2000. Alice: lessons learned from building a 3D system for novices. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 2000), ACM Press, New York, 486-493.

[9] Bowman, D. A., Wineman, J., Hodges, L.F., and Allison, D. 1998. “Designing Animal Habitats Within an Immersive VE”. IEEE Computer Graphics and Applications, 18(5), 9-13.

[10] Kitamura, Y., Fukatsu, S., Masaki, T., and Kishino, F. 1998. Intuitive control of “bird's eye” overview images for navigation in an enormous virtual environment. In Proceedings of the Symposium on Virtual Reality Software and Technology (VRST 1998), ACM Press, New York, 67-76.

[11] Chittaro, L., Gatla, V.K., and Venkataraman, S. 2005. A Navigation and Examination Aid for 3D Virtual Building. In Proceedings of IFIP International Conference on Human-Computer Interaction (INTERACT 2005), Springer Verlag, Berlin, 1014-1017.

[12] Stoev, S.L., and Schmalstieg, D. 2002. Application and taxonomy of through-the-lens techniques. In Proceedings of the Symposium on Virtual Reality Software and Technology (VRST 2002), ACM Press, New York, 57-64.

[13] Hirose, K., Ogawa, T., Kiyokawa, K., and Takemura, H. 2006. Interactive Reconfiguration Technique of Reference Frame Hierarchy in the Multi-viewport Interface. In Proceedings of IEEE 3D User Interfaces (3D UI 2006), IEEE Computer Society Press, Los Alamitos, 73-80.

[14] Tory, M., Moller, T., Atkins, M.S., and Kirkpatrick, A.E. 2004. Combining 2D and 3D views for orientation and relative position tasks. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 2004), ACM Press, New York, 73-80.

[15] Andrienko, G., Roberts, J.C., and Weaver, C. (eds.) 2007. Proceedings of the International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV 2007), IEEE Computer Society Press, Los Alamitos.

[16] Schafer, W.A., and Bowman, D.A. 2005. Integrating 2D and 3D views for spatial collaboration. In Proceedings of the International Conference on Supporting Group Work (GROUP 2005), ACM Press, New York, 41-50.

[17] Darken, R. P., Cevik, H. 1999. Map Usage in Virtual Environments: Orientation Issues. In Proceedings of IEEE Virtual Reality conference (VR 1999), IEEE Computer Society Press, Los Alamitos, 133-140.