Evaluation of the Command and Control Cube

Jérôme Grosjean*  Jean-Marie Burkhardt  Sabine Coquillart  Paul Richard
i3D INRIA  INRIA - Universite Paris V  i3D INRIA  ISTIA - Innovation
Jerome.Grosjean@inria.fr

Abstract

Application control in virtual environments (VE) is still an open field of research. The Command and Control Cube (C³) developed by Grosjean et al. is a quick access menu for the VE configuration called workbench (a large screen displaying stereoscopic images). The C³ presents two modes, one with the graphical display of the cubic structure associated to the C³ and a blind mode for expert users, with no feedback. In this paper we conduct formal tests of the C³ under four different conditions: the visual mode with the graphical display, the blind mode with no feedback and two additional conditions enhancing the expert blind mode: a tactile mode with the tactile feedback of a Cyberglove™ and a sound mode with a standard audio device. Results show that the addition of sound and tactile feedback is more disturbing to the users than the blind mode. The visual mode performs the best although the blind mode achieves some promising results.

1. Introduction

Virtual environment (VE) configurations have greatly enhanced the way of interacting with 3D scenes. New interaction techniques have emerged to provide users with tools adapted to the 3D world. However, most of the work has been concentrated on 3D navigation and object manipulation techniques. Application control is still an open problem in VE. On the one hand, WIMP (Windows, Icons, Menus and Pointing) paradigm commonly adopted in Desktop 2D applications cannot efficiently be implemented in VE, due to the intrinsically different nature of the interactions. On the other hand, it is necessary to increase systematic evaluations of each newly proposed techniques in order to give a strong empirical basis for developing 3D Interface standards.

To address the first issue, Grosjean et al. developed a new paradigm, the C³(command and Control Cube)[9] for the VE configuration called Responsive Workbench (hereafter referred to as workbench)[10, 11]. This menu presents a cubic shaped visual interface with a two-level selection process according to the user’s expertise. In the learning stage, the selection is done by moving a cursor inside the cubic shape. In this mode, the novice user is helped in the task by the visual feedback of the cube and the cursor. In the expert mode, users accustomed to the C³ are supposed to trigger the same commands by doing the selection without looking at the displayed graphics.

This paper deals mostly with the second issue by reporting on an experiment with 24 users performing with the C³ in 4 test conditions. Two of the four conditions are the usual modes, i.e. the learning mode (visual feedback) and the expert mode (no visual feedback, i.e. blind mode for items selection in the cube). Two additional conditions are built by coupling the expert mode with the assistance of another modality, i.e. audio-augmented expert mode (no visual feedback, audio feedback) and tactile-augmented expert mode (no visual feedback, tactile feedback).

The first section will briefly describe the C³. After reviewing the previous work, the experiment and the results are presented. In the final part, implications and further work are discussed.

2. Related Work

Application control inside virtual environments is still an open problem. Applications developed for these environments require an intuitive interface, adapted to the 3D world, to call functions or change the states of variables. At the SIGGRAPH’2001 course on 3D user interface design[1], Ernst Kruijff proposes a categorization for current control techniques influenced by the description of non-conventional control techniques by MacMillan et al[17]. He divides the approaches into graphical menus, voice commands, gestural interaction and tool.

In this group of graphical menus, Kruijff distinguishes three different categories: converted 2D-menus, hand-oriented menus and 3D widgets.
2.1. Converted 2D Menus

On traditional workstations, a desktop environment with a 2D screen and manipulated by a keyboard and a mouse, the WIMP interface (Windows, Icons, Menu and Pointing) is now acknowledged as the most popular graphical interface. A first spontaneous approach has been to transpose classical 2D pull-down menus to VE. To that end they are represented as flat 3D objects floating inside the virtual space. The main advantage advocating for this approach is that it provides the user with interaction techniques that remain familiar.

However, selection inside a flat floating menu is not convenient. Picking an item inside a menu is conceptually a 1D task. Both additional dimensions, especially depth, require unnecessary precision on the part of the user for the selection task [4, 8, 14].

Using a virtual ray controlled by the hand as a pointer tool is one method to diminish this difficulty[18]. Another approach, based on the virtual tricorder device[21], suggests using “2D anchored menus”. The 2D menu is displayed at the position of the device held in the hand. Moves and selection inside the menu are done by pressing the mouse buttons. This method does not require precision, but the selection remains a slow process.

2.2. Hand-oriented Menus

Kruijff divides Hand-oriented Menus into two sub-categories: 1DOF menus, where the items are put on a circular object[13][20] and menus stored at a body-relative position, where the selection of items is dependent on the relative position of the hand. The $C^3$ presents some similarities with the first category, but uses all dimension of the space around the user’s hand, instead of just one. The second approach exploits proprioceptive feedback[15] 1. It provides strong advantages for direct manipulation (excellent control of one’s own hand), physical mnemonics (finding body-centered objects) and gestural actions (recall of actions). The $C^3$ makes an extensive use of proprioception too, by arranging the menu item around the hand of the user.

Another approach[7] suggests materializing the plane of selection by holding a flat physical transparent palette in the non-dominant hand. The graphical menu is displayed directly onto the transparent physical surface (thanks to a tracker placed on the palette). The dominant hand is then used to pick items by touching the surface. This method takes advantage of both the hand-held prop and body-centered aspects, but requires the user to hold a physical device at all time to access the menu.

Lindeman[14] conducted an empirical study on handheld windows comparing the absence or presence of passive-haptic feedback and bimanual versus unimanual interaction. A selection and a docking task were used for the test. His primary measures of performance were mean task completion time and accuracy. Trial time was significantly faster and accuracy better when haptic feedback was present.

In 2001 Bowman et al.[2] presented a novel menu, called TULIP, based on the pinching of fingers. Pinch Gloves$^T^M$ are used to detect the pinching of two fingers while a head-mounted display (HMD) allows the display of menu items on top of each finger. Bowman et al. carried out an evaluation of TULIP, comparing the menu to the other well-known VE menu types: floating 2D menus and the pen and tablet-based menu. Users with some prior VE knowledge performed quicker the menu selection task in the last trials with TULIP than with the floating menus, but not with the tablet-based one. However, the subjects did not reach asymptotic performance levels in the thirty trials.

2.3. 3D Widgets

Conner et al.[6] describe widgets as the combination of geometry and behaviour. System control functionalities are directly put onto specific objects of the 3D world. Widgets could be seen as a highly context sensitive “menu”.

2.4. Evaluation Frameworks

There are still few works on application control inside VE. Studies do not always provide formal thorough testings apart from collecting user’s impressions. Some researchers try to give a formal framework to test performances of new interaction techniques based on the multiplicity of factors (user types, input/output devices, interaction techniques, application types, task contexts, etc.). Poupyrev et al.[16] proposed a general framework for manipulation techniques in VE. They focused on basic direct manipulation tasks, like selection and positioning tasks. This set of basic tasks are intended to cover most of the manipulation scenarios, so they can provide general useful results for immersive manipulation techniques. Bowman et al. presented in 1999 a similar testbed evaluation framework, based on a taxonomy of lower-level tasks and the study of performance metrics[3]. The selection process of the $C^3$ is based on a simple spatial selection task that is evaluated in this paper.

3. Description of the $C^3$

The $C^3$ is a quick-access menu system developed by Grosjean et al. for a workbench configuration. Taking advantage of the three dimensions of space for the se-
A workbench is a table-like device where the top of the table is a screen. Two images are rear-projected at high frequency on the screen to create a stereoscopic display. The two images are filtered by shuttleglasses so that the left eye receives the image calculated for its point of view and likewise for the right one. The user perceives the 3D models as if they were floating above the table. A second vertical screen can be added to enlarge the field of view.

The $C^3$ proposes an approach similar to hotkeys in the 2D world (workstation with a keyboard) to rapidly issue a limited set of commands to an application running on the workbench. The $C^3$ is based on the idea of marking menus for 2D screens. In marking menus, a pie menu pops up around the mouse pointer after a short delay (less than one second). A novice user lets this delay elapse and moves in the right direction to select an item inside the pie. This method of selection (choosing a direction instead of moving inside a linear list) has been proven to be more efficient[5]. An expert user who knows where the options are situated inside the pie should move fast enough to make blind selections. The $C^3$ extends this concept in 3D for VE. A spherical yellow pointer (hereafter called sphere) appears in the center of a cubic structure whenever a user pinches the thumb and the index finger. The device used to captured the pinching action is a modified mouse, where the buttons have been moved on the top of the user’s fingers. The cubic structure is a volume divided into 3x3x3 smaller cubes, called “slots”. The sphere starts in the center slot and follows the movement of the user’s hand, which is spatially tracked with an electromagnetic tracker. The sphere is blocked by the $C^3$’s frontiers and thus cannot leave the whole volume of the 27 slots.

The user manipulates the $C^3$ with the non-dominant hand, in order to leave the dominant hand free to work with other devices on the main task of the application without having to deal with menu operations. When the user pinches his fingers the $C^3$ appears in front of the hand. The button is sustained while the hand moves in space. Then the sphere pointer matches the hand movements. When the user releases the button, the position of the sphere inside the cubic structure determines the option that has been triggered. One menu item is associated to each slot. The $C^3$ may have up to 26 options attached, since the central slot is reserved for the cancel action, i.e. exiting the menu without choosing an option. Icons are displayed on top of each of the small cubes to identify the function associated to the slots. For visibility purposes only one horizontal plane of slots is visible at once, the plane containing the sphere. The vertical movements of the sphere change the current visible plane.

4. Experimental Evaluation of the $C^3$

The main objective of this study was to evaluate the $C^3$’s approach, and the support it provides to the users for the task of repeatedly selecting items within numerous options in VE. To address this issue, an experiment has been designed to collect and analyse objective performances of subjects using the $C^3$. The task was the 3D equivalent of the task of selecting items through 2D menus. Due to the spatial specificity of the selection process with the $C^3$, we were interested in investigating the effect of the item’s localization within the cube on the performances (e.g. speed and accuracy). Furthermore, we wanted to experiment the effect of the four following conditions of interaction. Two of the four conditions are the envisionned usual modes, i.e. the learning mode (visual feedback), the expert mode (no visual feedback, i.e. blind mode for items selection in the cube). Two additional conditions are constructed by coupling the expert mode with the assistance of another sensorial mode, i.e. audio-augmented expert mode (no visual feedback, audio feedback) and tactile-augmented expert mode (no visual feedback, tactile feedback). These latter conditions were supposed to ease the selection process by providing additional informations in comparison with the “blind” expert mode.

In the audio-augmented expert mode (see Figure 1), a short beep is emitted each time the pointer crosses a frontier between two slots. The tactile-augmented expert mode (see Figure 2) works identically by emitting a small vibration on the user’s thumb and index finger when crossing the frontiers. A Cybertouch™ (a glove with vibrators on the tip of the fingers) was used to produce the vibrations. Since we wanted to test the $C^3$ in its standard use, i.e. selection with the non-dominant hand, the Cybertouch™ was a left-hand glove.

4.1 Subjects

The user population consisted of 24 subjects (19 males and 5 females). Four of them were familiar with the workbench. None had previous experience with the $C^3$ menu.
The subjects were right-handed.

4.2 Task

The task was to sequentially select each of the 26 $C^3$'s slots. The order of trial was randomly assigned.

One issue was to decide how we could tell the users to select the functions inside the $C^3$ without having to train them first with a specific set of functions and their position inside the cube. The name of each function could not be used as input for non-expert users.

Each slot being a combination of simple space directions (left, right, front, back, top, bottom) an initial idea was to tell the users the correct combination of directions for the slot every time (e.g. the front upper left option). However, this solution was rejected because it could favor one selection strategy, in this case : decomposing the hand movements, over the other ones, like moving diagonally. It was undesirable that the input described the path to the correct slot.

We finally chose to display in front of the user a visual representation of the 27 slots and to highlight the current option to be selected (see Figure 3). The three planes of slots were placed a little farther from each other than in the $C^3$ so that all the slots were visible. Visually showing and highlighting the correct slot inside a cubic display avoids an additional cognition process. Users know instantly where the target slot lies without receiving any information on a specific path to the target. For instance they are free to decompose their movement along the cube axes or move diagonally.

4.3 Procedure

After a short presentation of the workbench, each subject started reading a small description of the $C^3$ and how to manipulate it. Then the subject performs some training trials in the 4 different modes, beginning with the visual feedback for a better understanding and ending with a mode different from the first of the test. The experiment begins after these first trials.

The four modes were tested in a different order for each of the 24 (=4!) subjects. For each experimental condition the user wore the Cybertouch™, the pinching device and the shutterglasses. The target cube appeared in front of the users and highlighted in white the slot to be selected. Then the system waited for a selection with the $C^3$, and 2 seconds later gave the result achieved. A wrongly selected slot was highlighted in red while a correctly selected slot was changed to green. Another two seconds later the system highlighted the next slot in white and everything went on.

All slots were tested twice, except the center slot which was reserved for the cancel action and not examined here. Three random selections were added at the beginning of each experimental condition to avoid any haste that could occur at the start up. The user was told about their existence. A break was done after each of these $2*26+3=55$ series of selections. Some questions were then asked about how the users felt with the $C^3$ and also at the end of the whole test.

4.4 Collected Data

Data stored were the discrete flow of the tracker positions, the ordered list of the slots which were proposed to the users and their effective selection, and finally the different times of selections (when the signal appeared, when the user activated the $C^3$, when the user ended the selection). The speed of selection was directly computed from this data.

5. Results and Discussion

The data were corrupted for one subject. The following analysis has been done with 23 subjects. An analysis of Variance (ANOVA) with repeated measure has been performed on the two variables speed and accuracy for the different conditions : no feedback (blind mode), with visual feedback, with sound feedback and with tactile feedback.
The speed is significantly affected by the condition (F(3,66) = 4.42, p < 0.0068). Post tests (HSD Tukey) indicate that the subjects quickly performed the selection in the visual condition (m = 1.0s) compared to the other conditions (blind : m = 1.2s ; sound : m = 1.3s ; tactile : m = 1.3s). This result shows a predominant benefit of the visual feedback on the subjects, which could be explained by the fact they are all novice users. They still need the support of the graphical display of the C³ to feel comfortable with the selections and make quicker confident movements.

The accuracy is also significantly affected by the condition(F(3,66)= 10,616, p < 0.0001). The average percentage of correctly selected slots is 92.8% with visual feedback, 87.0% in the blind mode, 83.5% with sound feedback and 84.7% with tactile feedback. One point of interest in this study was to see whether the use of the non-dominant hand for the manipulation and the absence of graphical elements was too demanding for the users. Since it was a first contact with the workbench, with a 3D menu system and especially this new selection paradigm for most users, these first performances are encouraging for the blind mode, although the non visual additional feedbacks (sound and tactile) did not help the users. The performances in the visual mode, although close to 100%, are not perfect. One explanation is the constraint of speed which was required during the test.

The height of a slot inside the C³ (position) has also a significant effect(F(2,44) = 134,058, p < 0.0001). Each accuracy value represents the number of correct selections performed in the nine cells of a plane (thus with a maximum of 18). The number of accurate selections is higher in the visual condition (m = 16,219), followed by the blind condition (m = 15,261), the tactile condition (m = 14,909) and the sound condition (m = 14,674). Post tests (HSD Tuckey) show that the visual condition is significantly different from all the other conditions. The highest score is obtained in the central plane (m = 17,352), and then in the upper plane (m = 15,587). The lowest score corresponds to the plane under the hand (m = 12,859). Movements at the same height for the arm were easily executed. One explanation could be that movements in a horizontal plane around the starting rest position of the hand are less demanding for the arm than movements which necessitate changing the height of the hand. It should also be noticed, that the presence of the horizontal screen of the workbench which hinders large movements in the lower direction. In this plane, movement requiring the user to stretch his arm the most are the least successful, while the ones closer to the body, requiring a less demanding flexing of the arm, have slightly better results.

The best results are achieved in the middle plane, with slightly lower performances when the hand has to move close to the user’s body and slightly better performances when moving away. In this case again, it seems that the users are more comfortable with movements necessitating a reasonable flexing of the arm.

6. Conclusion and Future Work

The C³ presented by Grosjean et al. is a new 3D oriented application control paradigm for VEs. This paper focuses on formal testings for this paradigm. The accuracy and speed of users in a selection task has been evaluated for each of the subjects and for each slot of the C³ menu. The performance when the C³ is not displayed and without any type of alternative feedback is encouraging for novice users, especially considering the fact the non-dominant hand is used for the manipulation in this paradigm.

Two additional modes have been tested in terms of assistance for the “blind” expert mode. In the first one a short
sound effect is played when the pointer crosses the frontiers between slots. In the second, a Cybertouch™ is used to provide a tactile feedback, applied under the same conditions. These additional feedbacks were judged more disturbing than helpful and obtained surprisingly worse results, even than the single blind mode. Users have troubles in fact distinguishing the one signal frontier crossing from the two or three signals crossings, when the diagonal movements are perfect or near perfect and the signals very close in time from one another.

A significant effect of the position of the slots inside the $C^3$ is obtained. The different heights of the planes of slots perform differently, with the central plane being better than the upper one, and the lower plane being the worst of the three. The lower performances of the lower plane may be partially explained by the awkward presence of the horizontal screen of the workbench under the user’s hand, preventing some freedom of movement. In most cases, the position requiring the least bending of the arm seems to yield the best performances, while movements too close or too far from the body appear to be less accurate. This chart of performances for each position gives some indication to application designers on where the fast or frequent access options should be placed inside the $C^3$ and thus where the less time sensitive option could be best placed.

Future work includes evaluating the learning process of this new application control paradigm. The selection time for the blind mode, while slightly above the mean time obtained for pull-down menus[19], should be further tested with expert users. The blind mode intended for trained users could then be compared to the performances of its keyboard counterpart, hotkeys.

**References**


