

Evaluating Advanced Interaction Techniques for Navigating Google Earth

Dubois, E., Truillet, Ph., Bach, C.

IRIT – University of Toulouse 3

118, route de Narbonne

F – 31 062 Toulouse Cedex 9

+33 5 61 55 74 05

{Emmanuel.Dubois ; Philippe.Truillet ; Bach}@irit.fr

ABSTRACT

This paper presents the design and comparison of a mouse-based interaction technique (hereafter IT) and two advanced IT, used in public spaces to support navigation in a 3D space. The comparison is based on a composite evaluation, including performance and satisfaction aspects. These preliminary results demonstrate that the use of mixed IT in a public space do not result in more differences among user than a mouse-based IT. It also highlights the fact that performance and satisfaction have to be considered simultaneously since they appear to be two complementary aspects of an evaluation, especially in public space environment, where the performance is no longer the only dimension to consider.

Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems]: Artificial, augmented and virtual realities. H.5.2 [User Interfaces]: Evaluation methodology, Interaction styles, User-centred design.

General Terms

Design, Experimentation, Human Factors.

Keywords

User-centred design, experimentation, augmented reality, mixed interactive techniques, user-testing.

1. INTRODUCTION

Recent developments have pushed human-computer interfaces beyond the traditional mouse/keyboard configuration, into novel and multiple display systems, multi-sensory and multi-modal interfaces and virtual and augmented reality. The use of these advances in the design of input and output devices is intended to increase the usability of the interaction techniques. Commensurate with this revolution, there has been a great deal of research into 1/ information visualisation including metaphors, 3D representation, views such as perspective walls or fisheye views and 2/ adapted interaction techniques (hereafter, IT) such as large screens, dedicated tools or mixed and augmented systems.

Visualisation exploits the user's visual acuity, cognitive abilities, expertise and experience. For instance, large displays can help to build situation awareness through a common understanding of the information presented and to facilitate coordination. To allow users' intuitive as well as collaborative exploration, interaction capabilities need to be improved so that access to the data and associated features become apparent.

However IT developed to take advantage of these considerations are rarely compared with previous results and their effectiveness is seldom quantified by user studies. To better understand the impact of these multiple factors, it is becoming crucial to develop new design guidelines and metrics for usability evaluation of these interactive systems and environments. Much research has focused on the development of visualisation toolkits [10], new dedicated devices ([2], [3], [11]) and new interaction techniques [13] on large displays. Nevertheless, the situation of use often includes critical environments such as medicine, military command posts or air traffic control in which the user has primarily to be high-performing.

In our work we are examining a slightly different kind of context: advanced interactive techniques in public spaces. Indeed, our work is part of a larger project that aims at developing mixed interaction techniques, such as tangible UI or Augmented and Mixed Reality UI, in the context of museums. The goal is to transform knowledge into an interactive experience that carries knowledge and involves the user deeply. Evaluating user's performance remains important because the user must not be slowed by the interaction technique. However, other aspects such as discovery, pleasure, integration, satisfaction and robustness are also relevant when evaluating interactive systems in such public spaces.

The goal of the present work is thus to include performance and users' satisfaction aspects in a comparison of different interaction techniques in a public space context of use. Our first hypothesis is that these two aspects are not necessarily correlated: similarly to ergonomic criteria, designer will always have to establish a compromise between these two design options. We consider that reducing the gulf of execution [8] is extremely important in such contexts. Our second hypothesis is that providing interaction techniques that clearly separate the available commands will have a positive impact on their use. Finally, we also believe that reducing the input articulatory distance by providing interaction techniques that require user's action in direct correlation to the command to apply, will help the user to rapidly adopt skills for manipulating the technique.

In the rest of the paper, we first present the concrete application we used the basis of this experiment. We then introduce the interactive techniques we developed to investigate the three hypotheses mentioned above. We finally detail the settings of our experiment and discuss the first major results.

© Dubois, E., Truillet, P., Bach, C., 2007

Published by the British Computer Society

Volume 2 Proceedings of the 21st BCS HCI Group
Conference

HCI 2007, 3-7 September 2007, Lancaster University, UK
Devina Ramduny-Ellis & Dorothy Rachovides (Editors)

2. MOUSE-BASED AND ADVANCED INTERACTION TECHNIQUES

Google Earth (GE) [4] is a free application that supports navigation, bookmarks and search onto satellite pictures of the Earth. Pictures, presented from a bird-eye view perpendicular to the surface of the globe, are mapped on a sphere representing the Earth. The point of view can be modified, thus providing the user with a pseudo 3D view on these images. In this experiment we only considered navigation features that can be activated without any use of the Google Earth menus, icons and navigation tool provided in the upper right corner of the standard application. Furthermore, semi-automated features such as “double-click” or “double-click and slide” are implemented but were excluded from our experiment settings. The next sections present how the available commands are accessible with the mouse and the two advanced interaction techniques that we developed.

2.1 GE manipulation with a mouse

A first set of commands is used to support translations of the displayed image of the earth. For example, in order to visualise a region of the globe situated on the left of the current screen, one must press the left mouse button, move the mouse to the right and release the mouse button: this results in a rotation of the globe from left to right (see Fig. 1). Likewise, translations of the mouse to the left, top or bottom result into displaying globe areas placed on the right, bottom or top of the current view, respectively. In addition, the use of the mouse wheel enables the user to modify the altitude of the bird-eye view: this corresponds to the definition of the zoom level on the images.

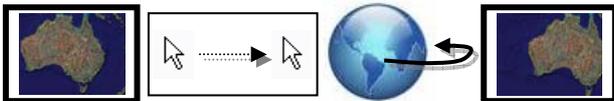


Figure 1: Effect on Google Earth of the mouse translation to the right while pressing the left mouse button.

The second set of commands considered in this experiment results in rotations of the displayed images. If the mouse cursor is in the upper half part of the displayed globe area, pressing the wheel mouse and then moving the mouse cursor horizontally to the left (resp. right), results in a counter-clockwise (resp. clockwise) rotation of the globe area as shown in fig. 2: this corresponds to a modification of the orientation of the North direction, and this behaviour is inverted if the mouse cursor is in the lower half part of the screen.

Finally pressing the scroll wheel and then vertically moving it down (resp. up), results in a diminution (resp. increase) of the angle between the point of view and the globe surface tangent (see fig.2): this corresponds to a modification of the viewpoint.

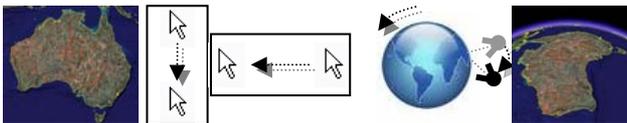


Figure 2: Effect on Google Earth of a combination of the horizontal and vertical translations of the mouse while pressing the scroll wheel.

2.2 GE-Stick

The first interaction technique we developed is called the Google-Earth Stick. This interaction technique is interesting for two reasons: first, it clearly separates the different available commands (one per sensor) and second, user’s actions that have

to be performed are similar to the result of the command (turn the potentiometer to turn the earth, push the slider to zoom out, etc.) which should reduce the gulf of the execution (hypothesis 2).

It consists of a prop held in the user’s hand and a board representing a compass rose. To perform the 4 translations, the user has to bring the prop close to one of the 4 directions represented on the compass rose (top, bottom, right and left). Bringing the prop close to one of the two areas present in the middle of this compass rose and tilting the prop up or down modifies the orientation of the point of view. Finally, there are two buttons on the prop: one can be turned with the thumb and forefinger to modify the orientation of the North axis and the second can be slid up or down to change the zoom. In a specific and predefined position, the “neutral zone”, these buttons have no effect.

This technique is based on the Phidget sensors [9]: the prop includes an RFID reader, a potentiometer and a slider; on the reverse side of the board, RFID tags are fixed to detect the position of the prop on the board.

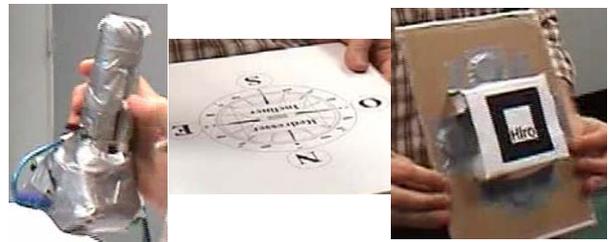


Figure 3: GE-Stick and the two sides of the board.

2.3 GE-Steering Board

The second interaction technique we developed is the Google-Earth Steering Board. This interaction technique is notable as, firstly, the user is no longer in direct contact with any device thus reducing the risk of damaging or stealing devices, especially in public contexts. Secondly, user’s actions (move the board to the left) are directly applied to one of the domain concept: the viewpoint on the image (move the point of view to the left) which should reduce the input articulatory distance.

The same board as used with the GE-Stick now represents the position of the point of view on the images. Moving the board results in moving the point of view accordingly; a neutral zone also exists in which only rotations can be triggered.

Technically, the GE-Steering Board involves video-based tracking software [1] to localise the position of the reverse side of the board in the space. As illustrated in fig. 4 a camera for the detection of the board is positioned in front of the user.

3. EXPERIMENTAL SETTINGS

To investigate these techniques, we conducted an experiment comparing mouse interaction with GE-Stick and GE-Steering Board in navigation tasks.

3.1 Users and Material

14 users were involved in this experiment, 8 males and 6 females (29.6 years old, SD = 7.3). All of them are researchers in Computing Science, which means they are familiar with mouse use. The version of Google Earth we used is 4.0.2416 and it was displayed on a screen 2.1 m wide, 1.5 m high (2.56 m in diagonal): the picture was retro-projected on the screen. Users stood in front of the screen, at approximately 2 m. A table was

placed between them and the screen at 1.9 m. An area was defined on the table to represent the zone in which to manipulate the mouse; it also represented a vertical projection, onto the table, of the “neutral zone” defined for the GE-Steering Board: this area was 0.3 x 0.25 m.

A camcorder placed on the same side of the screen as the user, captured data displayed on the screen and the totality of the user’s interaction (see fig. 4). On the other side of the screen, hidden from the user, two observers measured the duration of each task accomplished by the user and took notes about misuse of the interaction techniques.



Figure 4: Picture of the settings in front of the screen

3.2 Procedure

A pre-study usability evaluation session was performed on the three techniques and was based both on expert usability inspection and analyses of elementary interaction. These pre-experiment usability sessions were based on previous lessons learnt: cleaning up an ongoing evaluated IT avoids problems than can jeopardize user- testing and facilitates interpretation of results.

Each user was involved in a 3 phase process: training, measurement and interview. During the training phase, the goal was to teach to the user how to perform the ten different Google Earth commands involved in the experiment: six translations, including the zoom in and out, and 4 rotations. Users were informed that semi-automated zooming and moving with mouse are prohibited by experiment settings. The user received concise instruction from an experimenter before trying to achieve each command: the instruction was just composed of a description of the Google Earth command, e.g. “move toward the top of the screen”, without any information explaining the link between the command and the interaction technique: the user had to discover by himself how to use the interaction technique. Before teaching the next command, users had no time limit and were asked to confirm whenever they thought they perfectly understood and controlled how to trigger the command. Observers behind the screen, recorded the duration of the different task taught to the user, with the different techniques. Each user had to go through this training with the mouse first: this technique constitutes our control technique; users were then trained on the GE-Stick and GE-Steering Board in a counterbalanced order. Finally, using the mouse only, cities involved in the measurement phase were shown to the user (Paris, New York, Nouméa).

During the second phase, users were informed that the time to perform the following tasks was measured. This measurement phase was based on a predefined scenario. The scenario was

made of seven steps involving the ten Google Earth commands previously taught to the user: for example, users were all starting from Paris and the first step asked the user to “reach Liberty Island at an altitude of 400 m”. Each step of the scenario was stated by the experimenter who explicitly mentioned when to start carrying out each step. Each user had to perform the scenario three times, with the three different interaction techniques, in the same order taught in the training phase.

During the last phase, users were interviewed about their experience with the application. The goal of this semi-guided questionnaire was to identify the preferred technique, the most complex one, and the best and worse aspects of each technique. These interviews were not recorded and the experimenter leading the interview was in charge of taking notes of the answers.

4. ASSESSMENT RESULTS

4.1 Satisfaction analysis

A semi-guided interview was used to determine for each user, the preferred IT, their feelings about the discovery process, the 3 strongest and weakest points of each techniques, the most efficient and the most constraining techniques, and finally the physical workload. In this paper, we focus on the user’s preferred IT.

To identify preferred IT, participants were asked to rank ITs by preference order. A proportional score (range 0 to 20) was then extracted by summing the scores given by the users to each IT. According to this analysis the preferred IT is the GE-Steering Board (12.14) followed by the mouse (10.71) and GE-Stick (7.14). Preliminary interview analyses indicate that users prefer the GE-Steering Board because it allows a good level of presence and a feeling of omnipotence. Preference for the mouse seemed to be based on familiarity arising from everyday uses (habit).

This result tends to confirm our third hypothesis, that having to perform a user’s action similarly to the behaviour of the command has a positive impact on the use of the IT.

4.2 Performance analysis

In order to complement the previous analyses, we conducted analyses on the duration of use of each IT (sum of the durations of the seven steps). This paper focuses on descriptive performance results based on means (M) and standard deviations (SD).

M and SD of the duration of use (in minutes) over the 14 users have been calculated for the mouse (M=5’19; SD=2’31), the GE-Stick (M=6’48; SD=0’53) and the GE-Steering Board (M=7’23; SD=2’13). It thus clearly appears that the performance order, based on the duration of use, is completely different from the preferred IT order. Significance of the difference has not been studied so far.

This result confirms our first hypothesis that performance and satisfaction are not necessarily correlated. It also illustrates the necessity of a composite evaluation (multi methods and domains) of advanced interaction techniques as mentioned by [7]. Beyond this first comparison between performance and satisfaction, a composite evaluation approach is also important to refine evaluation results: for example, the analysis of the SD identifies the GE-Stick as the most stable inter-users IT which confirm our second hypothesis that a strong differentiation of each command leads to more consistent inter-users performance; however this IT is the least preferred. These different nuances clearly show that a simple performance analysis is still

insufficient to evaluate the quality of an advanced interaction technique.

Another performance result arises from the comparison of the SD of each IT. The SD represents the level of variability between users' performance when interacting with an IT: the smaller the variability, the more consistent the IT in terms of stability of use among different users. Unexpectedly, it appears that the SD of the two advanced ITs are equal (GE-Steering Board) or smaller (GE-Stick) than the SD of the mouse (control group). Since stability of use is one of the major factors of transferability of an IT to public spaces, because it positively affects consistency of use, this interesting result suggests that advanced interaction techniques are worth being further investigated and developed in such a context.

When considering the results in terms of M of duration of use, the better performance is accomplished with the mouse. However, we noticed during the experiment that, with mouse use, the speed of the image translations was directly correlated with the speed of the user's movement of the mouse, while with the two other ITs, the speed of the image translations is constant, even when user's movement are quicker or larger. Technical constraints caused these differences and further analysis based on the video-records will be conducted to define a ratio to correct the values obtained with the mouse on one hand and the two other ITs on the other hand, or to modify the link to the software application.

Additional preliminary analyses of the duration of each step of the scenario confirm results of pre-study usability evaluation sessions: some IT seems to be more appropriate for specific navigation commands. This result suggests it would be useful to perform a more detailed analysis of the duration of each step.

Finally, we have informally observed inter-user performance differences: this third axis will also be further investigated, especially from the point of view of a possible effect of IT presentation order. Indeed, we have identified that some users had difficulty understanding one experimental instruction during the measurement phase whereas they well understood the same instruction during the training phase. Each of these users had learnt to manipulate the GE-Stick before the GE-Steering Board. For now, we think that this situation could be a consequence of the cognitive workload due to GE-Stick training as mentioned in similar context by [12] for this kind of fragmented IT. Another hypothesis is that it could be a simple consequence of a misunderstood instruction. Finally, this may have been the result of a problem with the recruitment. Indeed, despite all the precautions that we have taken with the recruitment process, we observed, for one user, an important performance deviation (i.e. outside of $M + 2 SD$) that can be explained by user disabilities with orientation and 3D. These difficulties to control sample consistency are mentioned by other authors and led them to avoid some problematic user's profiles [5] or to use an elaborate specific recruitment process based on tests measuring participant's abilities to interact with Virtual Environments [6]. These points led us to take this subject out of the group.

5. CONCLUSION

This paper shows the main and most interesting results and questions of a first iteration of a larger user-centered evaluation process of advanced interaction techniques dedicated to public spaces. The first lesson learnt from this work is that that it is impossible to be satisfied by a performance evaluation on its own when assessing such advanced interaction techniques. A composite evaluation process is required: indeed reducing an evaluation to performance considerations may result in exclud-

ing other solutions better suited to the user's experience and satisfaction. Using composite evaluation will be useful to encourage the investigation of technical improvements of IT that are not the most efficient but that are the most appreciated by the users.

The composite evaluation conducted in our experiment and with our users sample, also shows that our prototypes could well be transferred to public spaces, because they appeared to be of better quality than mouse on various dimensions such as inter-users stability (GE-Stick) and our satisfaction measure (GE-Steering).

This paper only reports preliminary results of this experiment and further analyses of our data will address different hypotheses or open questions extracted from this iteration. We intend to focus on the identification of good and bad points of each technique according to the different steps of the scenario, the study of possible IT presentation order effect on IT learnability or cognitive workload (e.g. could be a warm topic in a multipoint exhibition context), the analysis of the recruitment process of participants (e.g. pre-test users' abilities and knowledge about Virtual Environment interaction required or not), and the improvement of the technical and software realisation of our advanced IT.

Further experiments will also be conducted to compare our advanced IT with other interaction techniques such as Wii-mote and to evaluate the usability of these techniques with other applications.

6. REFERENCES

- [1] AR Toolkit web site: www.hitl.washington.edu/artoolkit
- [2] Atlas Gloves, atlasgloves.org
- [3] Frölich B., Plate J., The cubic mouse: a new device for three-dimensional input, CHI'2000, The Hague (Netherlands), p. 526-531
- [4] Google Earth web site: <http://earth.google.com/intl/en/>
- [5] Green, C. S. and Bavelier, D., Effect of action video games on the spatial distribution of visuospatial attention. *Journal of experimental psychology: Human perception and performance*, 32, pp. 1465-1478, 2006.
- [6] Griffiths, G., Sharples, S., & Wilson, J. R., Performance of new participants in virtual environments: The Nottingham tool for assessment of interaction in virtual environments (NAIVE). *International Journal of Human-Computer Studies*, 64, 240-250., 2006.
- [7] Hartson, H. R., Andre, T. S. and Williges, R. C., Criteria for evaluating usability evaluation methods. *International Journal of HCI*, 13, p. 373-410, 2001.
- [8] Norman D., *The Design of Everyday Things*, MIT Press, 4th printing, 2001
- [9] Phidgets web site: <http://www.phidgets.com/>
- [10] KitWare, <http://www.kitware.com/>
- [11] SpaceNavigator, www.3dconnexion.com/products/3a1d.php
- [12] Stanney, K.M, Mourant, R. R. and Kennedy, R.S., Human Factors issues in Virtual Environments : a review of the literature. *Presence : Teleoperators and Virtual Environments*, 7, pp. 327-351, 1998.
- [13] Tse E., Shen C., Greenberg S., Forlines C., Enabling Interaction with Single User Applications through Speech and Gestures on a Multi-User Tabletop, AVI '06, May 23-26, 2006, Venezia, Italy, pp. 336-34