A Study of Street-level Navigation Techniques in 3D Digital Cities on Mobile Touch Devices

Jacek Jankowski, Thomas Hulin, Martin Hachet

A Study of Street-level Navigation Techniques in 3D Digital Cities on Mobile Touch Devices

Jacek Jankowski* Thomas Hulin Martin Hachet†

Inria Bordeaux, France

ABSTRACT

To characterize currently most common interaction techniques for street-level navigation in 3D digital cities for mobile touch devices in terms of their efficiency and usability, we conducted a user study, where we compared target selection (Go-To), rate control (Joystick), position control, and stroke-based control navigation metaphors. The results suggest users performed best with the Go-To interaction technique. The subjective comments showed a preference of novices towards Go-To and expert users towards Joystick technique.

Index Terms: H.5.2 [Information Interfaces and Presentation (e.g. HCI)]: User Interfaces—Interaction Styles, Evaluation/Methodology;

1 INTRODUCTION

3D digital cities are becoming widely available. Numerous acquisition campaigns are taking place all around the world. Thanks to advances in computer graphics technology, the resulting high quality 3D content can be rendered in real time on commodity mobile devices, potentially through a network connection. This opens interesting possibilities for the emergence of new usages dedicated to the masses. For example, imagine that you are planing a trip in a city you do not know. Currently, to explore the environment around your hotel in advance to get familiar with the surroundings and discover restaurants or other point of interests, you could use 2D maps and systems that provide successive 360-degree panoramic views like Google Street View. We believe that in the near future, you will be able to choose to use 3D environments that will offer richer, more interactive content in which you could freely navigate.

One of the limitations of the current 3D digital cities is the difficulty for the users to interact with them in a straightforward way [11]. Indeed, the control of the virtual camera may be difficult for novice users, particularly when interaction takes place on a touchscreen. Standard interaction techniques that were designed for general purpose navigation tasks on desktop computers may not fit well the particular case of urban navigation on touch devices [14]. The goal of our work is to investigate this particular context by experimenting with the interaction techniques that appear as the best candidates to complete a locomotion task in a 3D virtual city. We identify the pros and cons of each approach thanks to a user study that we describe in the following sections.

2 IDENTIFYING THE BEST TECHNIQUES

Navigation refers to the process of getting around a virtual environment: it includes a motor component called travel, and a cognitive one, known as wayfinding [2]. Numerous navigation techniques have been developed to aid both expert as well as novice users when navigating through VEs [10]. In this section we summarize approaches that we belive are most suitable for street-level navigation in 3D digital cities for mobile touch devices.

*e-mail: jacek.jankowski@inria.fr

†e-mail:martin.hachet@inria.fr

2.1 Target Selection (Go-To)

One of the most popular approaches for navigating in 3D digital cities is based on (1) indicating a destination point where the camera should automatically fly to and (2) dragging to orient the camera. This technique, based on Mackinlay et al.'s *Point of Interest Logarithmic Flight* [12], and now also known as *"Go-To"*, is a standard functionality of current commercial products like GoogleEarth or Google StreetView. This approach, illustrated in Figure 1(a), will be one of the experimental condition of our evaluation.

2.2 Rate Control (Joystick)

Another approach for navigating in 3D cities is to control directly the speed and orientation of the camera following the *vehicle* metaphor [16], with a first order transfer function. *Flying* or *driving* techniques controlled by keyboards and mice are standard on desktop 3D viewers. These techniques are harder to control on mobile touch-devices [6]. For such devices, in particular when used for videogames, a virtual joystick is generally preferred to move the camera (see Figure 1(b)). The orientation of the viewpoint may be controlled either with another virtual joystick, or by sliding a finger on the touchscreen with a zero-order transfer function similar to the one used in the previous section. These configurations will also be tested in the user study.

2.3 Position Control (Slide-Grab)

Position control can also be an interesting alternative to move the camera viewpoint. In particular, the *click and grab* metaphor of 2D map viewers extends well to 3D globe viewers as soon as the scene is viewed from an exocentric point of view. In this case, the 3D point that has been picked remains under the cursor/finger taking benefit of the *through-the-lens* approach [5]. On the other hand, this approach may become unadapted when the user is navigating inside the city at a street level. Indeed, few pixels on the screen may corresponds to large distances in the 3D world, resulting in inaccurate movements. This inaccuracy is increased on touchscreens because of the fat finger problem. Consequently, for the user study, we implemented a version where the camera movements are mapped to the user's gestures in the screen space, as done in Drag'n Go [13], but where the location of the initial point has no effect on the trajectory (Figure 1(c)). Since preliminary tests showed that some subjects perceive this mapping as going *"in the opposite direction"*, we let the subjects choose their favorite mapping between sliding the finger up for i) "sliding the camera forward" (Slide) or for ii) "pushing the world away" (Grab). In our implementation, orienting the camera is performed with a dual-point input.

2.4 Stroke-based Control (Draw)

We have also developed and evaluated a constrained 3D navigation technique [4, 8] which applies street topology in order to limit the navigation space. Similarly to [15, 14], we assume that the road network is given as a graph with the set of nodes representing the terminal and intersection points of the streets and the set of edges connecting the nodes that represent the streets. Traveling in this mode consists of moving along streets and selecting from available streets at crossings with user-drawn gestures as commands for moving the camera. Contrary to the *path drawing* technique [9] where a path is drawn in the 3D space, the stroke-based gestures in our

(a) Target selection (b) Rate control (c) Position control (d) Stroke-based control

Figure 1: The four interaction techniques selected for our experiment.

case operate in the screen space (like in [7]) to avoid problems of inaccuracy as mentioned above.

With this technique, users do not need to control the camera all the time. Instead, they sketch movements on the touchscreen to indicate the camera how to move. To keep the technique simple, we only use basic gestures: vertical and horizontal gestures as well as a tap gesture. A vertical straight gesture towards the top (resp. bottom) of the screen starts an animation that moves the camera forward (resp. backward) along the current street. When the user tap the screen while traveling, the movement stops immediately.

When the motion is stopped, the user can turn the camera from street to street by inputing horizontal gestures. When the camera is located at a crossroad, right and left gestures result in orienting the viewpoint towards one of the connected streets. When the camera is stopped somewhere in a given street, these gestures imply a rotation of the camera to the opposite direction. Technically, this consists in looking for the adjacent nodes in the street network. The current street that has been selected and that faces the camera is highlighted with a blue line displayed on the ground, as illustrated in Figure 1(d). This provides a visual feedback to the users allowing them to know what is the current state of the system.

In addition to vertical and and horizontal gestures, we also provide the users with a *Go-To* functionality that is controlled with tap inputs, when the camera is stopped. This allows them to move everywhere on the ground, without being constrained by the street network. The target location is temporary connected to the street network, so the user can continue interacting with the stroke gestures described above (Figure 1(d)). The orientation of the camera is controlled with a dual-finger input, similarly to the techniques we have described previously. Note that drawing a stroke have an effect only if the view direction is almost oriented towards a crossroad, which can be easily perceived by the users thanks to the visual feedback provided by the blue line.

A pilot study was performed to find initial values of speed levels for all the techniques. Five participants were asked to choose the parameters they liked best; the results and comments were used to improve the interaction styles for the main study.

An optimal 3D user interface is always a compromise. All of the previously described techniques are different in terms of cognitive load and motor effort as well as efficiency and ease of use. To better understand these differences we designed this user study that we describe in the following.

3 Evaluation Setup

3.1 Participants

16 subjects (5 female) with normal or corrected-to-normal vision participated in the experiment. The participants ranged in age from 19 to 44 (m=28). 5 participants were students and 11 had higher education. All of the participants were familiar with map navigation; 8 of them had no or very little experience with 3D interaction (we will refer to them as to novice users); 7 had very little experience with touch interfaces. Subjects were given gifts worth 5€/$7 for their participation.

3.2 Apparatus and Stimuli

The experiment was conducted on the first generation of Google Nexus 7, a 7-inch tablet computer with a 1280x800 pixel multi-touch capacitive touchscreen display. The test application used for the evaluation was developed using Unity3D. We prepared a small (about 10km^2) virtual 3D city representing the most common real-world street layouts. Twelve paths of similar complexity, 4x1.5km for the training and 8x2.5km for the main part of the evaluation, were defined. For each route, five cars: a red bus, a red firetruck, a brown humvee, a black-and-white police car, and a yellow taxi were parked on the way in a random order.

3.3 Procedure, Tasks and Measurements

Each test session started with an introduction to the test application and the four interaction techniques - the users were educated on how to use evaluated metaphors; they could also select orientation technique for joystick and mapping for position control technique (see previous section). The introduction was followed by a training session (one practice tasks for each interface) to allow the subjects to get familiarized with the test procedure. After the subjects indicated that they were satisfied, we proceeded with the actual trials.

We decided to use two types of tasks (based on [1]) representing various conditions a user is likely to experience using a 3D map application while being in the street level:

- *Travel Task (Camera Movement)* - For the first task, we asked the participants to simply follow a predefined path marked with red arrows projected on streets (see Figure 2(b)). Once they arrived at the end of the path, they were asked to recall positions of the cars they have seen on the way (Figure 2(d)).
- *Travel & Look-Around Task (Camera Movement and Orientation* - This task was similar to the first task. Additionally, we asked the participants to *shoot* green balloons that were deployed on the way (15 balloons for each route). The positions of the balloons were clearly marked in the environment (see Figure 2(c)). To *shoot* a balloon the user had to orient the camera directly towards the target and wait for one second.

The subjects were asked to complete the assigned tasks *"as accurately and as fast as possible"*. They were also told that it was more important to solve the tasks correctly rather than to be quick. After being presented with all 8 tasks (4 UI * 2 tasks), the users were given a questionnaire and asked to directly compare the evaluated techniques. Each evaluation session lasted approximately 60 minutes. Presentation of variables was counterbalanced by means of Latin square design.

The study measured relative effectiveness of the interaction techniques by both examining travel component (time taken to finish tasks/balloon shooting accuracy) and examining cognitive component (distance between the placed car marker and the car's actual position - measure also used in [3]). Using the questionnaire (based on continuous Likert scales) we measured participants' subjective impressions on the interaction techniques.

(a) Layout of our 3D city (b) Travel task (c) Look-Around task (d) Cognitive task

Figure 2: Evaluation setup: (a) Our virtual 3D city consists of 5 parts representing the most common real-world street layouts; Figures (b), (c) and (d) present the evaluation tasks we used in the study. For more information refer to the supplementary video.

4 Results

We analyzed our results with analysis of variance (ANOVA). We modeled our experiment as a repeated-measures 4x2 design (Interaction Technique x Task). Bonferroni procedure was used for evaluating the pairwise comparisons.

4.1 Objective Results

As overall balloon shooting accuracy was very high (0.995%), we decided to let this measure apart. During the learning phase of the experiment, when selecting orientation style for the Joystick technique, four users (all of them experts) chose second joystick. As for the position control technique, five users preferred *slide* and the rest chose *grab* metaphor.

Analysis of the task completion time revealed significant main effect of interaction technique ($F(3, 45)=10.591$, $p<0.0001$). Post-hoc comparisons of means revealed that the Go-To condition resulted in the best overall task performance ($p<0.0001$). More precisely, executing tasks using the Go-To technique was about 29% faster than using second fastest Joystick technique (136s vs. 192s).

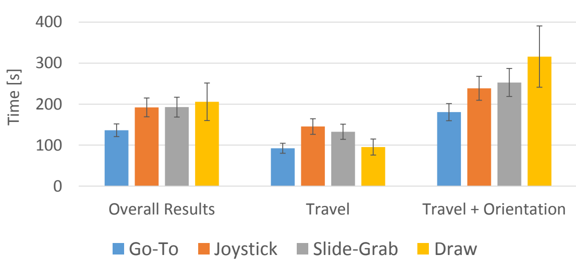


Figure 3: The completion time results (overall and for each task).

We also found a significant main effect of interaction between interaction technique and task type ($F(3, 45)=21.233$, $p<0.001$) on task completion time. Post-hoc comparisons of means revealed that executing tasks using the Joystick technique was significantly slower than using the Go-To and the Draw techniques for the travel task ($p<0.05$). For the travel + orientation task, the Go-To condition resulted in the best overall task performance, followed by the Joystick and Slide-Grab techniques, followed by the Draw technique ($p<0.0001$). Figure 3 illustrates the overall results of our experiment and the results for each task with respect to task completion time (error bars denote 0.95 confidence intervals).

Analysis of the recall distance did not reveal any significant main effects. Nonetheless, while looking at the effect of interaction technique ($F(3, 39)=2.4$, $p=0.081$), we observed a trend, where the subjects were more accurate in positioning cars in the recall phase for the Joystick and Slide-Drag navigation techniques. Figure 4 illustrates the overall results of our experiment and the results for each task with respect to recall distance.

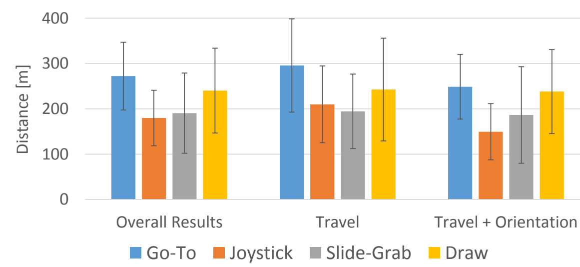


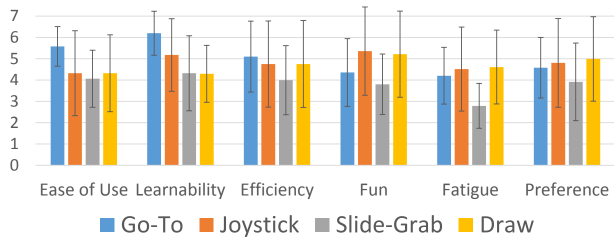
Figure 4: The results from the cognitive part of the study showing average distances between the placed car markers and the cars' actual positions. The smaller the distance, the better recall.

4.2 Subjective Results

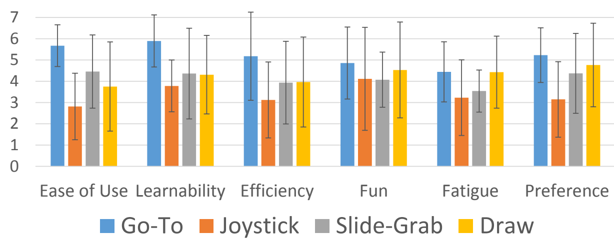
The average subject ratings with respect to ease of use (1:difficult/7:easy), learnability (1:difficult/7:easy), efficiency (completing the tasks fast and accurately was: 1:difficult/7:easy), fun (1:dull/7:fun), fatigue (1:tireing/7:not tireing) and overall preference are illustrated in Figure 5, together with standard deviations.

Analysis of the results revealed significant main effects in ease of use ($F(3, 48)=3.16$, $p<0.05$), learnability ($F(3, 48)=6.69$, $p<0.01$), and fatigue ($F(3, 48)=4.48$, $p<0.01$). Post-hoc comparisons of means revealed that the Go-To technique was perceived easier to use than the Slide-Grab technique ($p<0.05$). It was also perceived as easier to learn than the Slide-Grab and the Draw techniques ($p<0.01$). Moreover, the subjects perceived the Slide-Grab technique as more tiring than the Joystick and Draw techniques ($p<0.05$).

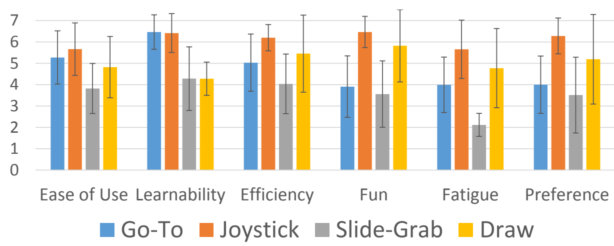
When considering only novice users, we found a main effect in ease of use ($F(3, 21)=4.39$, $p<0.05$): analysis revealed that the Go-To technique was easier to use than Joystick ($p<0.05$). As for the expert users, significant main effects of all variables were found ($p<0.05$). Most importantly, Joystick was perceived easier to use and more efficient than the Slide-Grab technique ($p<0.05$). The Go-To and Joystick techniques were perceived as easier to learn than the Slide-Grab and the Draw techniques ($p<0.01$). Using Joystick was more fun than using the Go-To and the Slide-Grab techniques ($p<0.01$), while Draw was more fun than the Slide-Grab technique ($p<0.05$). Finally, the expert users preferred Joystick over the Go-To ($p<0.05$) and the Slide-Grab ($p<0.01$) techniques.



(a) From all participants



(b) From novice users



(c) From expert users

Figure 5: Subjective results from the questionnaire.

5 Discussion

We can summarize the results from this comparative user study in the following main findings:

- Go-To navigation technique resulted in best overall travel task performance;
- The techniques requiring continuous input (the Joystick and Slide-Grab techniques) helped participants to be more accurate in the recall task;
- The subjective comments showed a preference of novice users towards Go-To and expert users towards Joystick.

Like most controlled user-based studies, this one had many limitations that restrict the generality of our findings: although we tested quite a complex city layout, we still managed to test only a small sample of all possible city topography types; furthermore, the study was performed on only one screen size and our task did not require the subjects to interact in a real world setting. Nevertheless, this study clearly allows us to better understand how users tend to navigate under different interaction paradigms.

The Go-To technique was ranked by the users as the easiest to use, probably thanks to the simple, one finger based mechanics of the interaction. It also resulted in best overall travel task performance, probably thanks to the logarithmic nature of this navigation technique (movement operates within controlled completion times so users can quickly travel long distances). On the other hand, the same logarithmic nature inhibited user's ability to gather information while traveling through a virtual environment. As a result, the Go-To technique performed worst for the cognitive task.

Joystick seems to be a good approach for users comfortable with 3D interaction. However, this technique is not adapted to novice users, who ranked it as most difficult to learn and use. Two of our participants became slightly motion sick while using it. This finding confirms the results from [14]. In any case, orienting the view through a zero order transfer function appears as better suited than moving the view with a second rate-controlled virtual joystick.

Regarding the position control technique, the majority of the subjects preferred the Grab metaphor, while others were more comfortable with Slide. This preference depends on cognitive interdependencies and, consequently, the choice of the technique should be let to the user.

Finally, what we have also learned is that the technique should be kept simple. For example, we hypothesized that the Draw technique would lead to a better performance. This happened to be true, but only when no precise positioning was required. Managing both long displacements along the streets and more precise movements, as it was the case for the balloon shooting task, appeared as too demanding for the users. Even when it was less efficient, some subjects preferred to input many successive target destinations with the Go-To technique rather than to mix two approaches, stroke gestures to go fast and target selection to refine.

To sum up, no technique clearly outperformed others for all navigation (travel and wayfinding) tasks. Consequently, the problem of navigation at a street level on touch devices remains open.

Acknowledgements

This work has been supported by the *Villes Transparentes* project.

References

[1] D. A. Bowman, D. Koller, and L. F. Hodges. A methodology for the evaluation of travel techniques for immersive virtual environments. *VR*, 3(2), 1998.

[2] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. 2004.

[3] N. Elmqvist, M. E. Tudoreanu, and P. Tsigas. Evaluating motion constraints for 3d wayfinding in immersive and desktop virtual environments. In *Proc. CHI'08*. ACM, 2008.

[4] T. A. Galyean. Guided navigation of virtual environments. In *Proc. SI3D'95*. ACM, 1995.

[5] M. Gleicher and A. Witkin. Through-the-lens camera control. In *Proc. SIGGRAPH'92*. ACM, 1992.

[6] M. Hachet and A. Kulik. Elastic control for navigation tasks on pen-based handheld computers. In *Proc. 3DUI'08*. IEEE, 2008.

[7] B. Hagedorn and J. Döllner. Sketch-based navigation in 3d virtual environments. In *Proc. SG'08*, 2008.

[8] A. J. Hanson and E. A. Wernert. Constrained 3d navigation with 2d controllers. In *Proc. VIS'97*. IEEE, 1997.

[9] T. Igarashi, R. Kadobayashi, K. Mase, and H. Tanaka. Path drawing for 3d walkthrough. In *Proc. UIST'98*. ACM, 1998.

[10] J. Jankowski and M. Hachet. A survey of interaction techniques for interactive 3d environments. In *Proc. EG'13 STAR Reports*, 2013.

[11] A. M. MacEachren and M.-J. Kraak. Research challenges in geovisualization. *CaGIS*, 28(1), 2001.

[12] J. D. Mackinlay, S. K. Card, and G. G. Robertson. Rapid controlled movement through a virtual 3d workspace. In *Proc. SIGGRAPH'90*. ACM, 1990.

[13] C. Moerman, C. P. M. Marchal, Damien, and L. Grisoni. Drag'n Go: Simple and Fast Navigation in Virtual Environment. In *Proc. 3DUI'12*. IEEE, 2012.

[14] A. Nurminen and A. Oulasvirta. Designing interactions for navigation in 3d mobile maps. In *Map-based Mobile Services: Design, Interaction and Usability,*. Springer, 2008.

[15] T. Ropinski, F. Steinicke, and K. Hinrichs. A constrained road-based vr navigation technique for travelling in 3d city models. In *Proc. ICAT'05*, 2005.

[16] C. Ware and S. Osborne. Exploration and virtual camera control in virtual three dimensional environments. In *Proc. SI3D'90*. ACM, 1990.