Poster: Guided Tour Creation in Immersive Virtual Environments

S. Pick[1*] A. Bönsch[1] I. Tedjo-Palczynski[1,2†] B. Hentschel[1] T. Kuhlen[1]

[1]Virtual Reality Group, RWTH Aachen University, Germany [2]Kisters AG, Germany

Abstract

Guided tours have been found to be a good approach to introducing users to previously unknown virtual environments and to allowing them access to relevant points of interest. Two important tasks during the creation of guided tours are the definition of views onto relevant information and their arrangement into an order in which they are to be visited. To allow a maximum of flexibility an interactive approach to these tasks is desirable. To this end, we present and evaluate two approaches to the mentioned interaction tasks in this paper. The first approach is a hybrid 2D/3D interaction metaphor in which a tracked tablet PC is used as a virtual digital camera that allows to specify and order views onto the scene. The second one is a purely 3D version of the first one, which does not require a tablet PC. Both approaches were compared in an initial user study, whose results indicate a superiority of the 3D over the hybrid approach.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/Methodology I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques

1 Introduction

Virtual Reality (VR) systems are used to interactively explore complex virtual environments (VEs), e.g. architectural models or numerical simulation results. Such scenes contain certain information that is often not obviously accessible to novice users, leading to two major problems: (1) Search Problem—Where can relevant information be found? (2) Route Planning Problem—How can relevant information be efficiently accessed? These problems can be addressed by using *guided tours*, which automatically guide users along a collection of paths from one view onto the scene to another, each of which provides visual access to a certain point of interest (POI). In order to define meaningful guided tours, it is important to determine what views are to be visited and in which order.

So far, many systems to create guided tours have been proposed. Beckhaus et al. [2] introduced the *CubicalPath* system that uses potential fields to guide users. *Way-Finder* by Andujar et al. [1] determines paths along which the amount of shown information is maximized. A similar approach is presented by Elmqvist et al. [3]. Hsu et al. [5] present a semi-automated camera motion design system that allows to manually specify views by various techniques. The *Navidget* by Hachet et al. [4] allows to specify views via 2D input on a tablet PC. It was adapted to immersive VEs (IVEs) [6].

While existing systems yield qualitatively good guided tours, they are often fully automated and do not support user-specified views [1, 3] or do not discuss interaction in IVEs [2, 5]. Also, the ordering of views is often not considered at all. To this end, we present two interaction techniques, a hybrid 2D/3D and a purely 3D one, that allow to manually specify and order views. A preliminary user study was conducted to initially evaluate our techniques.

*{pick|boensch|hentschel|kuhlen}@vr.rwth-aachen.de
†irene.tedjo@kisters.de

Figure 1: View definition (top) and ordering (bottom) with the hybrid (left column) and the 3D-only approach (right column).

2 View Definition (VD)

Both approaches use tracking data to define a view via a position and an orientation. For each view, a human-readable representation of it is stored in form of an image for later use during view ordering.

Our first approach is a hybrid interaction technique that uses a tracked tablet PC. It constitutes a virtual version of a digital camera as we surmise that it provides an intuitive metaphor to defining views. The tracking data is directly used to define a view and also render images of it to the tablet PC's display as feedback (see Fig. 1, top left). A button on the display—similar to a photo camera trigger—is used to store the view. Our second approach can be compared to a head-mounted action camera, as head tracking data is used to define a view. Based on camera finders, a semi-transparent rectangle provides feedback on the view that is currently being specified (see Fig. 1, top right). By pressing a button on a 6-Degree-of-Freedom (6-DoF) input device, the view is stored.

3 View Ordering (VO)

We use a matrix-like interface based on common photo browsers to enable view ordering (see Fig. 1, bottom). The advantage is a clear structure and sufficient space to present a number of images in a decent size at the same time. With the hybrid approach, the task is performed entirely on the 2D display of the tablet PC. Using the 3D approach, the matrix is displayed in the IVE and interaction is realized using raycasting with a 6-DoF input device. In the beginning of the ordering process, the matrix shows a set of numbered entries, each representing one potential view. By assigning an image to the entry with number *n* will make it the *n*-th view on the guided tour. The assignment is performed by first selecting an entry, after which a list of all stored images is shown—also using a matrix arrangement. By selecting an image, the associated view is assigned to the chosen entry. To add more views to the tour, the task is repeated. The tour can be modified by overwriting previous choices.

4 Preliminary Evaluation

We conducted a quantitative and qualitative within-subject user study to perform an initial evaluation of our techniques.

4.1 Experimental Design

A 5-sided CAVE was used for the study (3.6m $\times$ 2.7m $\times$ 2.7m $[w \times d \times h]$ at 1600×1200 pixels rectangular and 1200×1200 pixels square resolution). A 6-DoF A.R.T. Flystick was used for interaction. The hybrid interface was realized using a Samsung Q1-900 tablet PC (780g, 7" touch screen with 800×480 pixels). A model of a trade show booth was used as scene. Several objects in it had been annotated with short text labels to use them as POIs (see Fig. 1 top). The annotations eliminated the necessity for expert knowledge of the scene and allowed for a better comparability of subjects' results. Navigation was realized using the pointing travel technique. The user's motion was constrained to the only floor of the scene. For the 3D approach, we used the aforementioned Flystick. For the hybrid approach, a tracking marker was attached to the user's hand and a wireless button to the same hand's index finger.

Subjects had to perform one test run for each interaction technique. Each test run comprised two sub-tasks, one for VD and VO each. In the VD sub-task six views had to be reproduced which were shown to the user via a printout. It showed the image for each of the six target views, i.e. a labeled object. Users only had to roughly reproduce the view shown in the image and not the exact same view. After all six views had been defined, they had to be arranged into a pre-defined order in the VO sub-task. The target order was again provided via a printout. The initial interaction technique was alternated between subjects who could briefly familiarize themselves with the techniques at the beginning of a test run.

During a test run, the view definition and ordering times were measured. The view definition time was measured from the moment the user activated the VD technique to when a view was stored. To factor out naivgation as much as possible, users had to first navigate to a POI and were only then allowed to activate the VD technique. While it was active, correctional movement was only allowed by physical walking. Users were told to create only one view per POI. The view ordering time was measured from when the VO technique was activated to when a finish button was pressed. Users were asked to fill out a questionnaire after finishing both test runs in which their preferences were captured using 5-point Likert scales. Preference questions were accompanied by follow-up questions asking for the reasons behind choices.

4.2 Results

Overall, 22 subjects (avg. age 25.4 years, 2 female) participated of which 19 were considered for evaluation due to corrupted log files.

The average task execution time for the VD sub-task was 10.03s (std.dev. 8.27s) for the hybrid approach, while it was 7.08s (std.dev. 5.13s) for the 3D approach. For the VO sub-task the times were 43.93s (std.dev. 19.14s) and 46.64s (std.dev. 23.57s) respectively. We applied a two-tailed, paired Student's T-test to the measured task execution times. The execution times for the VD sub-task were found to be significantly different, favoring the 3D interaction approach ($t = -3.0, p = .008$). However, execution times did not significantly differ for the VO sub-task ($t = .565, p = .579$).

The quantitative results were supported by qualitative findings. Users were asked whether they prefered one approach over the other using a 5-point Likert scale ranging from 1 (prefer hybrid) to 5 (prefer 3D). The question was asked 3 times: once for each sub-task and the overall task. We performed a non-directional, one-sample Wilcoxon Signed-Rank Test on the answers to check for significant differences. In case of the VD sub-task a significant preference for the 3D approach could be determined ($z = 2.364$, $p < .02$). However, no sigificant preference was found considering the VO sub-task ($z = 1.764, p > .05$). A significant preference for the 3D approach was found regarding the overall task ($z = 2.26, p < .05$).

The follow-up questions helped us identify issues with both approaches. About two thirds of the users (12 during VD, 10 during VO) stated that they had to carry too many things. Users found the hybrid interface to be more familiar due to the camera metaphor, but perceived the size and weight of the tablet PC negatively ("too heavy" 13 times, "too big" 10 times). Five users even stated that the gear of the hybrid approach hampered their movement. The main issues of the 3D approach were its button layout (5 times, VD), a reduced input precision (4 times, VD), and reduced visual fidelity of the menus (4 times, VO).

Users were further asked to give reasons for why they preferred one approach over the other. In case of the VD sub-task, a better intuitiveness (6 times) and higher precision (4 times) were the main arguments for the hybrid approach. An easier handling (12 times), faster work speed (9 times), better intuitiveness and reduced physical strain (7 times each) were arguments for the 3D approach. Statements were similar regarding the VO sub-task: a better handling (3 times), higher precision, faster work speed and better visual fidelity (twice each) were given for the hybrid, and a better handling (7 times), faster work speed (5 times), and better intuitiveness (3 times) for the 3D approach. Overall, arguments for the hybrid approach were an increased feeling of control (4 times), less physical strain, less complexity and better learnability (twice each). For the 3D approach they were less physical strain (8 times), an increased feeling of control (7 times), and less complexity (6 times).

Only 2 participants stated that looking at two displays negatively affected immersion. However, 7 users felt the frequent changes between the two displays disrupted the workflow of the VD task.

5 Conclusion

We have presented two interaction approaches for the definition of views and their ordering for the inclusion in a guided tour. One approach is based on hybrid 2D/3D and the other on pure 3D interaction. A qualitative and quantitative within-subject user study was conducted to perform an initial evaluation. It indicated an advantage of the 3D approach over the hybrid one.

In the future we plan to optimize the approaches by addressing the issues that were identified during the user study. Additionally, we would like to compare our techniques to other view definition techniques in the context of guided tour creation.

Acknowledgements

This research was funded by the German Research Foundation DFG as part of the Cluster of Excellence "Integrative Production Technology for High-Wage Countries".

References

[1] C. Andújar, P. Vázquez, and M. Fairén. Way-Finder: Guided Tours Through Complex Walkthrough Models. *Computer Graphics Forum*, 23(3):499–508, 2004.

[2] S. Beckhaus, F. Ritter, and T. Strothotte. Guided Exploration with Dynamic Potential Fields: The CubicalPath System. *Computer Graphics Forum*, 20(4):201–210, 2001.

[3] N. Elmqvist, M. E. Tudoreanu, and P. Tsigas. Tour Generation for Exploration of 3D Virtual Environments. In *Proc. of the ACM Symposium on Virtual Reality Software and Technology*, pages 207–210, 2007.

[4] M. Hachet, F. Decle, S. Knödel, and P. Guitton. Navidget for Easy 3D Camera Positioning from 2D Inputs. In *Proc. of the IEEE Symposium on 3D User Interfaces*, pages 83–89, 2008.

[5] W.-H. Hsu, Y. Zhang, and K.-L. Ma. A Multi-Criteria Approach to Camera Motion Design for Volume Data Animation. *IEEE Trans. on Visualization and Computer Graphics*, 19(12):2792–2801, 2013.

[6] S. Knödel, M. Hachet, and P. Guitton. Navidget for immersive virtual environments. In *Proc. of the ACM Symposium on Virtual Reality Software and Technology*, pages 47–50, 2008.