Poster: Investigating Viewpoint Visualizations for Click & Go Navigation

Benjamin Nuernberger*   Steffen Gauglitz*   Tobias Höllerer*   Matthew Turk*

Department of Computer Science, University of California, Santa Barbara

ABSTRACT

We present an investigation of viewpoint visualizations for "Click & Go" 3D navigation interfaces based on a pre-populated set of viewpoints. These scenarios often occur in 3D navigation systems that are based on sets of photos and possibly an underlying 3D reconstruction. Given these photos (and the 3D reconstruction), how does one most effectively navigate through this environment? Existing systems often employ Click & Go interfaces which allow users to navigate with one click of the mouse or tap of the finger. In this work, we investigate viewpoint visualizations for such Click & Go interfaces, describing a preliminary user study and providing valuable insights into Click & Go and its viewpoint visualizations.

Index Terms: I.3.6 [Computing Methodologies]: Computer Graphics—Methodologies and Techniques; H.5.2 [Information Systems]: Information Interf. and Presentation—User Interfaces

1 INTRODUCTION

Many tools exist to reconstruct a physical environment in 3D based on imagery (e.g., [4]), and they continue to improve every day. To browse through such reconstructed 3D environments, two major challenges must be addressed: rendering the environment and a user interface to navigate it. In this work, we focus on the latter.

There has been much work in 3D navigation interfaces [2], including a class of interfaces which we term "Click & Go." Here, a single click takes the user to a new viewpoint; thus the new viewpoint has to be derived from an input with only 2 degrees of freedom (DoF). These interfaces are used in tools such as Google Street View, Microsoft's Photosynth (based on Photo Tourism [4]), and PhotoCloud [1]. Our interest in this class of interfaces stems from two reasons. First, Click & Go interfaces are ideally suited for *constrained* navigation, including the case when only a sparse, discrete set of viewpoints is available, which occurs in the aforementioned applications. Second, Click & Go is very simple, making it appealing even if the sparsity of the data is not a primary concern.

Since selecting a 6-DoF camera pose from a 2-DoF click location is an underconstrained problem, a Click & Go interface has to employ a *viewpoint selection* algorithm that includes certain assumptions about the user's intent. To communicate these assumptions to the user, most interfaces include a *visualization* of the viewpoint that will be chosen if the user clicks at the mouse cursor location.

Existing systems typically either visualize all currently available viewpoints (as done in PhotoCloud and the overview mode in Photo Tourism) or visualize only one viewpoint, chosen via the viewpoint selection algorithm based on the mouse cursor location. Furthermore, the visualization can provide information about the viewpoint (e.g., the frames in Photosynth and PhotoCloud) or the object of interest (e.g., the quadrilateral on building façades in Google Street View, which indicates the distance to the detected surface and its orientation).

The main contribution of this work is an exploratory user study of viewpoint visualizations for Click & Go interfaces with discrete

*{bnuernberger, sgauglitz, holl, mturk}@cs.ucsb.edu

(a) yellow: full size frame   (b) red: full size frame + origin

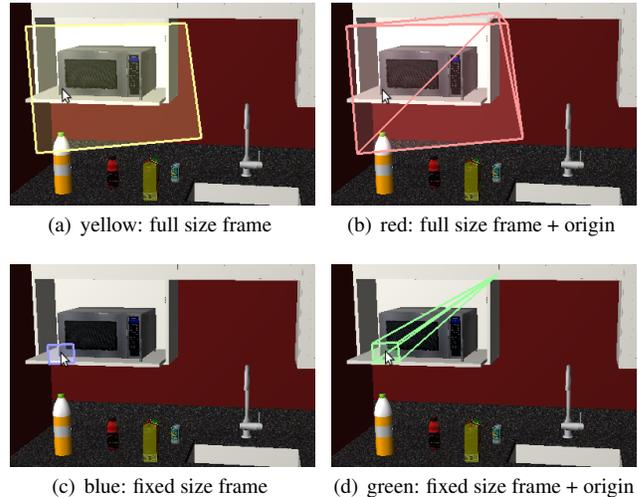(c) blue: fixed size frame   (d) green: fixed size frame + origin

Figure 1: Four of the five viewpoint visualization conditions compared in the main study; the fifth condition was no visualization.

viewpoints. We investigated how different visualizations are perceived by users and how much they influence the task completion time of an elementary primed search/viewpoint manipulation task.

2 USER STUDY

In total, 18 users participated in our study and were compensated for their time commitment of about one hour with a nominal amount. We tested various conditions on the first 8 users, including different types of tasks and a visualization similar to PhotoCloud's (i.e., all viewpoints are visualized simultaneously); however, this was determined to be too cluttered in our environment.

The main part of the study included 10 participants (18-30 years old; 5 female and 5 male; all passed a standard colorblindness test; 5 wore corrected lenses; and 2 had never previously used 3D software). Each task consisted of finding a specified piece of information (e.g., the brand name of an object, nutrition information on a food item, etc.) in a virtual kitchen scene. We intentionally removed the "search" aspect from the task by telling the user *where* the target object was located; thus, the only task was to navigate towards it.

We used a simple viewpoint selection algorithm (filter out viewpoints that cannot see the 3D world point P under the cursor; then choose the viewpoint whose optical axis is closest to P) and compared five different visualization conditions (Fig. 1). For easy reference and identification, we colored them in soft pastel colors; however, we explicitly told the participants to use the colors for this purpose only. The first four visualization conditions were chosen to reflect design choices from existing Click & Go systems, and the fifth condition was "no visualization." **Yellow** (Fig. 1(a)) and **red** (Fig. 1(b)) display a rectangle which indicates the camera's field of view at the depth of the moused-over object. **Blue** (Fig. 1(c)) and **green** (Fig. 1(d)) display a rectangle of fixed world size that is always centered around the mouse cursor and parallel to the camera's image plane. Red and green also indicate the camera's origin via a
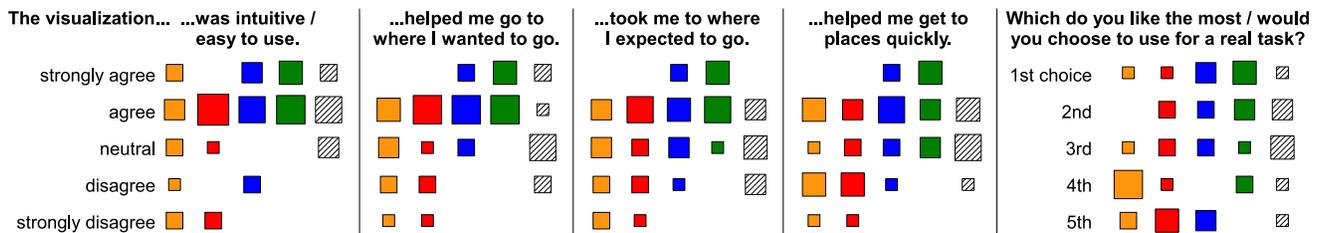
Figure 2: Aggregated responses from the post-study questionnaire. For each question, from left to right, the columns correspond to the yellow, red, blue, green, and "no visualization" conditions. Square sizes: 1 □ 2 ▫ 3 ▪ 4 ◾ 5 ◾ 6 ◼ 7 ◼ users.

pyramid wireframe. For the four visualizations, we darkened their color when their view's camera direction was opposite to the current camera view; this helped in overcoming ambiguity problems such as found with the Necker cube.

Users were able to rotate (i.e., look around) via dragging, but the only way to move the camera origin was to Click & Go in accordance with the visualization. Each Click & Go movement was animated by interpolating between the two poses over one second.

We tested all five visualization conditions in three different environments: (1) relatively few objects and 68 camera viewpoints spread out in the kitchen; (2) 93 camera viewpoints densely clustered around a moderate number of objects on a single countertop; (3) like (2), but with more objects, leading to a severely cluttered scene and the need to navigate around occlusions for most tasks.

We used a within-subjects design with task completion time as the dependent variable; the order of the conditions was fully balanced. For each visualization, each user completed two training tasks and five recorded tasks per environment, for a total of 10 training tasks and 75 recorded tasks (5 vis. × 3 env. × 5 tasks).

3 RESULTS

For each environment, task completion times were analyzed using a two-way repeated measures ANOVA. No significant effect of the visualization was found ($F_{(4,236)} = 1.85$, $F_{(4,236)} = 1.54$, $F_{(4,232)} = 0.49$, respectively, with $p > 0.05$ for all environments).

In a post-study questionnaire, users ranked the five visualizations and rated them on a 5-point Likert scale with respect to four questions (see Fig. 2). While the question "Which [visualization] do you like the most?" produced mixed results, users rated the two mouse-centered visualizations (blue and green) better than the two full-frame visualizations (yellow and red) in all four questions.

4 DISCUSSION

There are several possible explanations for why we did not observe any statistically significant effect of visualization on task completion time. First, we emphasize that this was a small-scale study with limited statistical power, designed to gain some insights (cf. Sec. 5) into viewpoint visualizations, rather than to prove the superiority of a particular visualization. While we tried to design the individual tasks to be of similar difficulty, we observed that they still varied appreciably in difficulty, which, while balanced across users and conditions as far as possible, increased the variance in the data.

Second, in very difficult cases such as cluttered close-ups, all of the visualizations suffered from certain drawbacks, such as being only partially visible, and thus the visual clutter and cognitive load of trying to identify the "best" view may not have paid off.

Third, to guarantee that each task was actually solvable, we made sure that there was at least one "good" viewpoint for each object of interest. Thus, in easy cases, our simple viewpoint selection algorithm was usually effective enough to select this "good" view even if the user just clicked on the object of interest without paying attention to the visualization at all. This may explain why the "no visualization" condition performed comparably to the other four con-

ditions (in some cases, "no visualization" produced even faster task completion times than the other four conditions).

We also note four remarks on the users' apparent preference for blue & green (from our observations and users' comments): (1) while yellow & red provide strictly more information than blue & green (i.e., the exact extent of the camera's field of view), this information was not always needed for the given task; (2) yellow & red sometimes take up a large portion of the screen, limiting their usefulness; (3) yellow & red may appear "jumpy," while blue & green appear more fluid since they are always dynamically centered around the mouse cursor; (4) finally, yellow & red make the user more aware of the discrete set of available cameras, which, unless specifically desired, may be considered to decrease presence.

5 CONCLUSION & FUTURE WORK

We presented an investigation into viewpoint visualizations for Click & Go interfaces. We discussed an exploratory user study on a target-oriented navigation task comparing five different visualizations. This leads to several insights and avenues for future work.

Although Click & Go is a simple interface, providing a good visualization for it is nontrivial. On the one hand, no visualization may even be necessary for simple environments with a good viewpoint selection algorithm, although it may be useful for discoverability of the Click & Go feature. On the other hand, geometrically complex scenes, especially close-up and cluttered environments, pose a significant challenge to the usefulness of current visualizations. In addition, users seem to prefer simple and fluid visualizations even if they provide less information. Thus, we suggest that a smart viewpoint selection algorithm, selecting predictable and possibly perceptually preferred [3] viewpoints, is very important and deserves more in-depth investigation.

Finally, future investigations should include different navigation tasks (besides target-oriented ones) and specific types of movement which we observed to be especially difficult to navigate via Click & Go, such as seeing an object from the other side (i.e., orbiting) and zooming in/out. Brivio et al. [1] also note similar difficulties.

REFERENCES

[1] P. Brivio, L. Benedetti, M. Tarini, F. Ponchio, P. Cignoni, and R. Scopigno. PhotoCloud: Interactive remote exploration of joint 2D and 3D datasets. *IEEE Comput. Graph. Appl.*, 33(2):86–96, c3, 2013.

[2] J. Jankowski and M. Hachet. A survey of interaction techniques for interactive 3D environments. *Eurographics – STAR*, pages 65–93, 2013.

[3] A. Secord, J. Lu, A. Finkelstein, M. Singh, and A. Nealen. Perceptual models of viewpoint preference. *ACM Trans. Graph.*, 30(5):109:1–109:12, 2011.

[4] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: exploring photo collections in 3D. *ACM Trans. Graph.*, 25(3):835–846, 2006.