The Effects of Visual Granularity on Indoor Spatial Learning Assisted by Mobile 3D Information Displays

Nicholas A. Giudice and Hengshan Li

Spatial Informatics Program: School of Computing and Information Science
University of Maine, Orono, Maine 04469, USA

giudice@spatial.maine.edu, hengshan.li@umit.maine.edu

Abstract. There is growing interest in improving indoor navigation using 3D spatial visualizations rendered on mobile devices. However, the level of information conveyed by these visualization interfaces in order to best support indoor spatial learning has been poorly studied. This experiment investigates how learning of multi-level virtual buildings assisted by mobile 3D displays rendered at different levels of visual granularity effect subsequent unaided navigation tasks. The visual granularity levels include: a high fidelity model, low fidelity model, wireframe model and sparse model. Results showed that using the sparse model during learning led to the most accurate and efficient overall pointing and navigation performance and that between-floor judgments were less accurate when assistance during learning was unavailable. These findings demonstrate that more information is not necessarily better and provide new insights into the optimal information content to be included in mobile 3D visualization interfaces supporting indoor spatial learning and cognitive map development.

Keywords: indoor navigation, 3D visualizations, mobile information displays, naïve realism, visual granularity, immersive virtual environments.

1 Introduction

Current advancements in the computational resources, memory capacity, and high-resolution display technologies available on mobile devices means that complex environmental visualizations are becoming a viable solution for real-time navigation systems. However, most existing navigation interfaces are limited to 2D representations and work exclusively outdoors. By contrast, our interest here is in designing indoor navigation systems based on 3D building visualizations. Considering that on average, people spend 87% of their time in indoor spaces [1] and since indoor built environments often are comprised of complex and confusing 3D spatial structures [2], providing access to a 3D visualization of the space (i.e., a ground-level egocentric map representation) is postulated as being advantageous and more realistic for supporting spatial learning and cognitive map development as compared to their traditional 2D analogs. Indeed, the efficacy of 3D visualizations and map representations for aiding

adfa, p. 1, 2011.

navigation through indoor environments is a topic of growing interest in both academic research [3-4] and for commercial applications, e.g. Google Maps and Nokia 3D indoor maps.

One practical question for these 3D visualization based navigation systems is how the realism of the 3D models affects human navigation performance? In outdoor environments, several authors from the geo-visualization and cartography communities have advocated the use of abstract rather than photorealistic 3D visualizations for more efficient inference making [5-6]. Empirical experiments addressing this issue support the view that users often have misplaced faith in realistic representations, termed "Naïve Realism" [7]. For example, people using spatial interfaces for naval applications prefer spatially realistic 3D icons of ships and planes on their displays vs. functional, symbolic icons. However, these realistic features were shown to actually decrease identification performance [7]. Similarly, users predicted they would need high-fidelity photorealistic 3D displays to find routes across outdoor terrain, whereas experimental results demonstrated that they actually performed the task better with lower fidelity displays [8]. Several studies have clearly shown that while photorealistic representations of maps appeal to users, they often have a negative impact on behavioral performance [9-10]. As was illustrated in Klippel et al. (2010), people trying to use Google street view for wayfinding purposes converged on a similar experience--that simply providing photorealism is not enough for accurate spatial learning and wayfinding [11].

However, few studies have been conducted to evaluate the effect of environmental realism of mobile interfaces supporting real time indoor navigation. In part, this is due to the lack of accurate indoor positioning for indoor environments and a dearth of real time indoor data models for use on mobile devices. Although relatively impoverished renderings are assumed to be as effective in aiding people's navigation through indoor spaces as photorealistic models, this assumption has not been extensively studied, although initial evidence has provided some empirical verification. For example, Kalia et al. (2008) found that richly rendered (photorealistic) indoor virtual models were not as efficient for spatial learning as a sparse model [12]. However, this study did not investigate different levels of visual granularity of 3D models, nor was it aimed at evaluating the efficacy of using a mobile navigation device to learn multi-level buildings, as is the goal here.

In this study, we experimentally evaluate four simulation fidelity conditions which manipulate the level of visual granularity of the environment which is provided to the user by a simulated mobile device during learning of virtual buildings. We aim to assess whether users' navigation performance after spatial learning with the mobile device differs as a function of the visual granularity of the interface, findings which will help specify the optimal information content to be used in future 3D displays for real-time indoor navigation systems.

The experiment was conducted using immersive virtual environments (VEs) rather than physical environments (PEs) as VEs best facilitate manipulation of building layout and information content, as well as tracking of movement behavior (see Fig. 1).

2 Methods

2.1 Participants

Twenty participants (10 female and 10 male, mean age=20.9, SD=2.0) were recruited from the University of Maine student body. All participants self-reported as having normal (or corrected to normal) vision. All gave informed consent and received monetary compensation for their time. There were three sessions for each subject, with each session lasting approximately one hour.

2.2 Materials and Apparatus

We used an SX111 HMD (NVIS, Inc), incorporating inertial tracking, a panoramic 111 degree field of view, and a high resolution 1260 x 1080 stereo display, which provides a highly immersive VR experience. Two Nintendo Wii remotes were used in the experiment. One was used by the experimenter to control the sequence of experimental phases, and the other was used by the participant to translate through the VE. Turning in the VE was done through physical body rotation.

Our environments were comprised of five two level buildings which were richly rendered in the VE. 3DS Max was used as the 3D modeling and rendering tool. The Vizard 3D rendering suite, by WorldViz Inc., was used as the VE platform supporting users' real-time navigation and recording their trajectory and test performance. As is illustrated in Fig. 1, two types of models were used in the experiment: virtual reality environment models and 3D visualization models. The former simulated the physical world in the VE and were made to be as photorealistic as possible in order to foster the experience of walking in the physical world. The latter included the 3D visualizations which were shown on the simulated mobile device during environmental learning.

Fig. 1. Simulated mobile device in the VE

Four levels of visualization granularity represent a natural progression of degraded surface detail for environmental rendering, while preserving building topology. Each model is depicted in Fig. 2. The high fidelity model was rendered with photorealistic texture, natural light, and full color (The Mental Ray rendering plug-in was used to generate the model. The low fidelity model used grey scale color to represent the building and there was no rendering of texture or photorealistic light. The wireframe model only rendered the lines at each edge. The sparse model was the simplest representation as it only contained the floor plan of each layout without walls and ceilings.

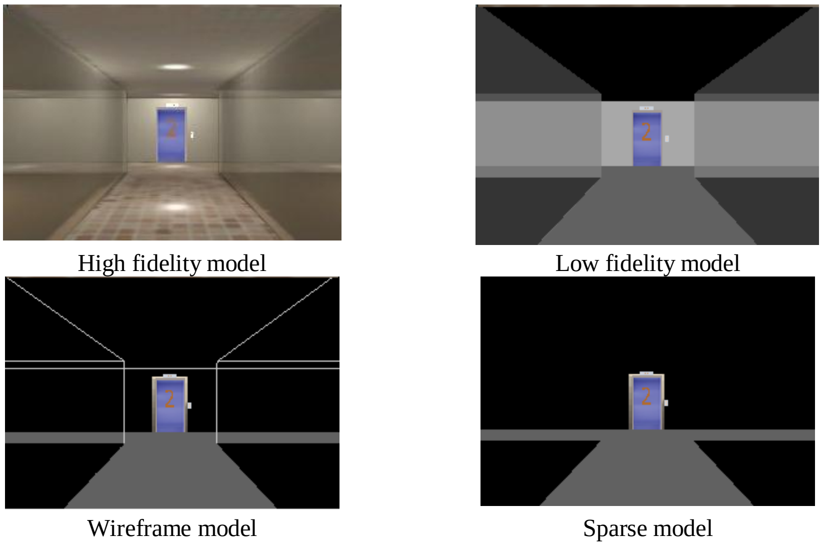

Fig. 2. Four visualization fidelity models as shown on the simulated mobile device

Each level of the building was based on a 3 x 3 matrix of hallways, as illustrated in Fig. 3. Each hallway was subdivided into two corridor segments. We deleted two segments from the twelve possible corridor segments in the generic environment to create our experimental layouts. This procedure ensured that all the layouts were well matched in terms of number of nodes, segments, and intersections.

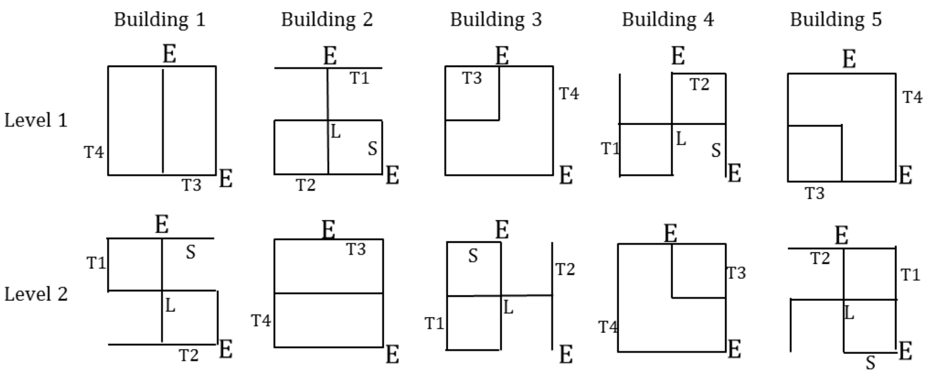

Fig. 3. Experimental building layouts

The two floors were connected by two elevators, which also served as salient landmarks for orientation in each of the experimental buildings ("E" represents the elevators in Fig. 3). From a top-down perspective, one elevator was always located at the top center and the other was located at the southeast corner. In Fig. 3, "L" represents the starting position during the learning period, which was located at the only 4-way intersection in the building. The starting position for the navigation tests, indicated by "S", was located near one of the two elevators to provide an orientation cue but was not visible from the starting learning point. There were two pictures on each floor which served as experimental targets, indicated by "T" in Fig. 3. Pictures were based on eight high imagery words: bottle, chair, clock, dog, fish, kite, table and tie. All routes between pictures were matched across building for route length and number of turns.

2.3 Procedure

A within subjects design was adopted, with twenty subjects running in all five visualization conditions. There were five phases in the experiment.

Phase 1: Practice. Subjects were familiarized with the apparatus and navigation behavior in the VE. All experimental tasks were explained and demonstrated before starting the experimental trials.

Phase 2: Route learning. In this task, participants learned the route to each picture with the assistance of the mobile device. From a north orientation at the learning start point, subjects were guided by arrows displayed on the mobile device to each target picture in each of the four visualization granularity conditions. After reaching the picture, which was hanging on the wall, they were asked to face the picture and remember its location. Subjects were then guided back along the same route to the learning start point and repeated the task for the next target. During the learning phase, the mobile device served as a navigation assistant as it provided increased visual access to the overall floor layout than was possible by simply looking around in the VE. In a fifth unaided control condition, the mobile device was not available during target learning; rather, guidance was done via arrows displayed on the ground. The outbound route for target learning was not necessarily the shortest route. Rather, we chose routes based on a trajectory that maximized environmental exposure. As such, if users looked around as they walked, as was the instruction, they could apprehend the entire building after traversal of the four learning routes. Overlap between routes was minimized to ensure no part of the building was over-learned.

Phase 3: Pointing criterion task. To test whether participants had successfully learned the four target locations from Phase 2 and could situate them in a globally coherent cognitive map of the building, they had to point to each target from the learning start point (target order was randomized by floor). The Phase 2 route learning and Phase 3 pointing task was done separately for each floor (floor order was counterbalanced). When making the pointing response, participants did not have access to the mobile device and no target was visible from the learning start point. Thus, accurate pointing required them to make Euclidean judgments from the learning start point to the target, with half of the targets located on a different floor. To meet criterion, participants needed to point to targets on each floor within a 15 degree tolerance. If they failed the first iteration, the Phase 2 learning and Phase 3 pointing tests proceeded until they either successfully met criterion or until they made four incorrect attempts. We recorded users' pointing time, angular error, and the number of iterations it took to pass the learning criterion test.

Phase 4: Re-exposure task. After the pointing test, participants once again walked from the start point to each of the four pictures (target order was randomized) with the assistance of the mobile 3D visualization interface in order to re-instantiate all targets in memory before starting Phase 5.

Phase 5: Unaided Navigation task. To perform this task, participants were positioned at the navigation start position as shown in Fig. 3. They were then given the name of one of the pictures and asked to navigate to it using the shortest route. This task was performed without assistance from the 3D visualization interface on the mobile device used during learning. The sequences of the pictures were pseudo-random to ensure two routes were within floor and two routes were between floors. Once they believe they had reached the picture, they pressed the button on the Wii mote to indicate its location and orientation. The sequence of pictures was counter balanced between conditions and participants. As subjects only traveled the route between the learning start point and each picture during the learning phase, determining the shortest route between pictures for this navigation task required accurate development and accessing of a "cognitive map" of the entire building. If the participant incorrectly indicated the picture's location or orientation, they were guided to its correct location and orientation before starting the next trial. This corrective measure was done to prevent the accumulation of error between trials. They were then asked to follow the same sequence of steps for the next target picture. This was done for four routes in total. Two dependent variables for the navigation task were analyzed. The first was navigation accuracy, based on whether subjects successfully indicated the correct location and orientation of the picture. The second was navigation efficiency, based on whether the shortest route was executed (e.g., shortest route length over traveled route length).

3 Results

3.1 Pointing Task

A repeated measures ANOVA on pointing angle error was run with visualization (5 levels: four granularity conditions and the unassisted control) and floor (2 levels: within and between floor target trials) as the within subjects factors. The within-between floor factor was significant, $F(1, 39) = 6.495$, $p < .015$, $\eta^2 = 0.143$, with the within floor absolute pointing error being 4.3 degrees lower than the between floor pointing error (98.3% of all pointing trials were within the 15 degree tolerance after 2 iterations). There was no significant main effect of pointing error as a function of visualization condition, $F(4, 156) = 1.138$, $p < .341$, $\eta^2 = 0.028$. However, subsequent pairwise comparisons showed that pointing error for between floor trials was significantly higher than for within floor trials with both the unaided (control) condition ($p<0.015$) and the low fidelity model ($p<0.027$).

Table 1. Mean pointing error (SE in parentheses) for within floor and between floors

	Unaided	High Fidelity	Low Fidelity	Wireframe	Sparse
pointing error within floor	5.3 (.9)	9.3 (2.4)	5.6 (1.1)	7.4 (1.4)	7.4 (1.3)
pointing error between floors	11.3 (2.9)	12.8 (3.3)	12.1 (2.9)	7.9 (1.7)	11.8 (3.6)

A repeated-measures ANOVA on pointing iteration trials was run with the same two within-subjects factors. A significant effect was observed for floor, with more iterations needed to pass criterion for between floor judgments (m = 1.24, SE = 0.05) than for within floor judgments (m = 1.11, SE = 0.023), F (1, 39) = 6.193, $p < .017$, $\eta^2 = 0.137$. There was no significant main effect of iteration as a function of visualization condition, F (4, 156) = 0.624, $p < .646$, $\eta^2 = 0.016$.

Table 2. Mean iteration (SE in parentheses) for within and between floor pointing judgments

	Unaided	High Fidelity	Low Fidelity	Wireframe	Sparse
iteration within floor	1.08 (.04)	1.18 (.06)	1.03 (.03)	1.13 (.05)	1.13 (.05)
iteration between floors	1.28 (.10)	1.28 (.09)	1.35 (.12)	1.13 (.05)	1.18 (.07)

A repeated-measures ANOVA for pointing time was run with the same within-subjects factors. Only floor was significant, F (1, 39) = 10.79, $p < .002$, $\eta^2 = 0.217$, with pointing time taking 3.3 seconds longer for the between floor judgments than for the within floor judgments. There was no significant main effect of pointing time as a function of visualization condition, F (4, 156) = 0.506, $p < .732$, $\eta^2 = 0.013$. However, there was a significant interaction between visualization level and floor, F (4, 156) =2.754, $p < .030$, $\eta^2 = 0.066$. Subsequent pairwise comparisons showed that pointing time for between floor trials was significantly longer than for within floor trials with both the control condition and the low fidelity model, each $p <0.005$.

Table 3. Mean pointing time (SE in parentheses) for within floor and between floor judgments

	Unaided	High Fidelity	Low Fidelity	Wireframe	Sparse
pointing time within floor	7.59 (.77)	8.93 (1.02)	6.79 (.68)	8.53 (.96)	8.80 (.82)
pointing time between floors	12.21 (1.69)	11.86 (1.97)	13.60 (2.16)	9.87 (1.23)	9.83 (1.26)

3.2 Unaided Navigation Task

A repeated-measures ANOVA for target localization accuracy during the navigation task was run with the same two within-subjects factors of visualization and floor. There was a significant main effect of target localization accuracy as a function of visualization condition, $F (4, 156) = 2.678$, $p < .034$, $\eta^2 = 0.064$, with localization accuracy after learning with the sparse model (m = 86%, SE=3.6%) being reliably higher than after using the low fidelity model (65%, SE=5.7%), p<0.001. The within-between floor factor was also significant, $F (1, 39) = 9.457$, $p < .004$, $\eta^2 = 0.195$, $\alpha=0.05$, with the navigation accuracy found for within floor performance (83%, SE=2.7%) being reliably higher than for between floor judgments (72%, SD=3.6%).

Table 4. Mean navigation accuracy (SE in parentheses) for within floor and between floors

	Unaided	High Fidelity	Low Fidelity	Wireframe	Sparse
navigation accuracy within floor	83% (6.1%)	83% (6.1%)	68% (7.5%)	90% (4.8%)	90% (4.8%)
navigation accuracy between floors	73% (7.1%)	70% (7.3%)	63% (7.8%)	70% (7.3%)	83% (6.1%)

A repeated-measures ANOVA for navigation efficiency was also run for the two within-subjects factors. There was a significant main effect of navigation efficiency as a function of visualization condition, $F (4, 156) = 3.192$, $p < .015$, $\eta^2 = 0.076$. Navigation efficiency with the sparse model (89%, SE=2.9%) was reliably better than the high fidelity model (73%, SD=4.9%) (p<0.008) and the low fidelity model (71%, SD=5.0%) (p<0.001). The within-between floor factor was not significant, $F (1, 39) = 1.641$, $p < .208$, $\eta^2 = 0.040$.

Table 5. Mean navigation efficiency (SE in parentheses) for within floor and between floors

	Unaided	High Fidelity	Low Fidelity	Wireframe	Sparse
navigation efficiency within floor	79% (5.4%)	76% (5.4%)	73% (6.9%)	87% (4.5%)	92% (3.1%)
navigation efficiency between floors	85% (5.4%)	71% (6.7%)	69% (6.6%)	73% (6.9%)	87% (4.9%)

4 DISCUSSION

The most important findings of this study are that using the sparse model to assist learning led to the highest unaided target to target localization accuracy and route efficiency performance. These results provide evidence that use of a sparse model of layout structure is better than both of the highest fidelity models for assisting environmental learning of complex buildings. These findings are consistent with, and extend, previous research regarding the evaluation of the realism of 2D maps [7-10]. One explanation is that participants need to extract picture and layout information from high fidelity 3D visualizations to encode the relative positions of these pictures as well as their positions in the building, whereas this information is more directly specified from the sparse model. This synthesis and extraction process may yield additional cognitive effort during learning which resulted in the increased navigation error and decreased efficiency for information-rich displays compared to the displays rendered with lower visual granularity.

We interpret the absence of significant differences for any of the five presentation conditions in the pointing task (pointing time, pointing iteration trials, and pointing error) as further demonstrating that adding realism to the 3D models during learning is neither necessary nor advantageous for extraction of Euclidean relations between targets and accurate cognitive map development.

As for the within-between floor analyses, our results are consistent with previous literature for multilevel indoor navigation [13]. Subjects took longer times to point, required more iterations to meet criterion, exhibited greater errors, and had lower navigation accuracy when pointing and navigating to targets located on different floors than when they were on the same floor. These results suggest that it is more difficult for people to maintain the spatial relation of objects between floors, likely made more difficult when inter-floor layouts are not congruent. Given the known challenges for integration of vertical knowledge in cognitive maps, future experiments will investigate new mobile visualization interfaces for integrating multi-floor buildings during indoor navigation. Importantly, the finding that the control condition showed reliably worse between floor pointing performance than the aided conditions (but for the low-fidelity model), indicates that having assistance during learning (e.g., providing better visual access), may improve knowledge of inter-floor relations. Indeed, we believe that performance in the control condition was likely elevated for all metrics in this experiment as our decision to maximize floor coverage during the route learning phase likely provided sufficient opportunity to apprehend global spatial relations, thereby reducing the inherent benefit afforded by the mobile devices to depict layout configuration. It is likely that performance in the unaided condition would have been significantly worse if we had used a more realistic route learning paradigm that emphasized minimum route length rather than breadth, and was done in buildings with greater topological complexity.

Taken together, these results provide compelling evidence that there is no reliable advantage of 3D information displays rendered at a high level of visual granularity on learning and navigation of buildings and that in many cases, the best performance is obtained using a sparsely rendered spatial model. To our knowledge, our results are the first empirical demonstration showing the advantage of using sparse models on portable mobile devices as supporting real-time learning and navigation of complex indoor buildings. As illustrated by Smallman et al. (2005), good display design is more than slavishly adhering to realism [7]. Our research extends the theory of naïve geography to use of 3D real time indoor maps and provides new evidence for the basic principle of these displays that graphics should not provide more information than is needed by the user [7]. Our results also provide an empirical foundation to help guide the development of more efficient visualization interfaces to be implemented on future indoor navigation systems.

5 ACKNOWLEDGMENTS

This research was supported by NSF grant number: IIS-0916219. We thank Bill Whalen and Chris Bennett for their comments on the manuscript and data analysis.

6 REFERENCES

1. Klepeis, N., Nelson, W., Ott, W., et al.: The national human activity pattern survey (nhaps): a resource for assessing exposure to environmental pollutants. Journal of Exposure Analysis and Environmental Epidemiology. Volume 11(3). (2001) 231-252
2. Giudice, N.A., Walton, L.A., Worboys, M.: The informatics of indoor and outdoor space: a research agenda. In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness. (2010) 47-53
3. Mulloni, A., Nadalutti, D., Chittaro, L.: Interactive walkthrough of large 3d models of buildings on mobile devices. In: Web3D '07: Proceedings of the twelfth international conference on 3D web technology. (2007) 17-25
4. Chittaro, L., Nadalutti, D.: Presenting evacuation instructions on mobile devices by means of location-aware 3D virtual environments. In Proceedings of the 10th international Conference on Human Computer Interaction with Mobile Devices and Services. (2008)
5. Haeberling, C.: Cartographic design principles for 3D visualization–A contribution to cartographic theory. In Proceedings of the 22nd ICA International Cartographic Conference, A Coruña, Spain. (2005)
6. Döllner, J.: Non-Photorealistic 3D Geovisualization. In Multimedia Cartography. Berlin, Germany. (2007) 229–240
7. Smallman, H. S., St. John, M.: Naïve realism: Misplaced faith in realistic displays. Ergonomics in Design, Volume 13. (2005) 14-19
8. Smallman, H. S., Cook, M. B., Manes, D. I., Cowen, M. B.: Naïve realism in terrain appreciation. In Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society. (2007) 1317 - 1321
9. Hegarty, M., Smallman, H. S., Stull, A.T.: Decoupling of intuitions and performance in the use of complex visual displays." Proceedings of the 30th Annual Conference of the Cognitive Science Society. Washington, DC, USA: Cognitive Science Society. (2008).
10. Hegarty, M., Smallman, H. S., Stull, A.T., Canham, M.S.: Naïve Cartography: How Intuitions about Display Configuration Can Hurt Performance. Cartographica: The International Journal for Geographic Information and Geovisualization Vol. 44(3). (2009) 171-186
11. Klippel, A., Freksa, C., Winter, S.: You-are-here maps in emergencies – the danger of getting lost. Journal of Spatial Science Volume 51. (2006) 117–131
12. Kalia, A.A., Legge, G. E., Giudice, N.A.: Learning building layouts with non-geometric visual information: The effects of visual impairment and age. Perception. Volume 37(11). (2008) 1677-1699
13. Holscher, C., Meilinger, T., Vrachliotis,G., Brosamle,M., Knauff, M.: Up the down staircase: Wayfinding strategies in multi-level buildings. Journal of Environmental Psychology. Volume 26. (2006) 284–299