Designing 3D Geographic Information for Navigation Using Google Glass

Kelvin Wong*[1] and Claire Ellul†[2]

[1]Department of Computer Science, University College London, London, UK, WC1E 6BT
[2]Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK, WC1E 6BT

January 05, 2015

Summary

No longer bound by traditional 2D physical representations, there is a steady shift towards three-dimensional (3D) data. Existing research recognises landmarks to be important navigational but specific geometric and semantic attributes in 3D have not been identified. This study offers a user-centred investigation into assessing of the saliency of environmental objects which facilitate pedestrian navigation. A novel real-world navigation experiment using Google Glass is carried out with fourteen participants. Results show geometric and semantic detail for navigation are most pertinent between 1.65 – 7.5m for buildings. Visual characteristics such as colour, shape and texture are more relevant than function and use.

KEYWORDS: 3D GIS, Navigation, Google Glass, Landmarks, User-Centred Design

1. Introduction

Navigation is an implicit requirement of our daily lives. In recent years, the way we traverse the world has been strongly impacted by the emergence of mobile navigational technologies, altering our perception of space and specifically, our navigational strategies. A complex and ever-evolving phenomena, the study of navigation has a long history in a diverse number of fields ranging from psychology to geography to computer science. Navigation experiments, however, remain predominantly in virtual and simulated environments. There is a lack of real world experiments examining human navigation behaviour, perhaps due to the lack of control and higher costs incurred. Where existing studies have found landmarks to be the most salient feature in an urban landscape for navigation (May et al., 2003), saliency of the 3D geometric and semantic attributes of these landmarks have not been identified. This study aims to address the above issues by exploring landmark saliency at a finer, intrinsic level through a real-world navigation experiment. The paper will outline the design of the user-centric experiment and the subsequent results from its first iteration. It will conclude with possible directions for future work.

2. Methodology

In order to capture the true dynamism of real navigation, fourteen subjects were asked to navigate in the real-world rather than in a computer-simulated environment. A novel approach using a pair of Google Glass was implemented to record the gaze and movement of the participants. The device is minimalist and light optical head-mounted display, allowing for unobtrusive and natural tracking. The experiment was carried out around UCL and the Bloomsbury, an area of predominantly residential and office spaces.

Participants were instructed to follow a specified route on the map provided. The selected route (1.518km) was designed to test the navigation strategies of the participants and passed through an area

* kelvin.wong.11@ucl.ac.uk
† c.ellul@ucl.ac.uk

without with major branding or obvious iconic buildings. By choosing a complex and diverse route, a significant navigation task demand was imposed on the participants and promoted the description of geometric and semantic features of navigational landmarks. Participants were told that the task was not to navigate the route most successfully or the fastest but were asked to produce a set of written instructions of the route for an unfamiliar traveller who had not been to the area before. This was to promote active thinking and to engage with the process of navigation. The participants were then fitted with Google Glass, presented with the route map and were allowed to navigate freely (Figure 1).

Figure 1 Partcipants wearing Google Glass and navigating

A Google Glass application was specifically designed and written for this experiment. Developed in Android Glass Development Kit Preview 4.4.2, the application runs on Google Glass XE 19.1 and was paired with a smartphone (iPhone 5S) in order to obtain locational information via GPS. The user does not directly interact with the Google Glass, but rather it passively records a first person video as well as tracking their gaze vector (orientation and pitch), location (latitude and longitude) and elapsed time. The video was recorded concurrently while the gaze vectors and location were logged every 0.125 seconds. Figure 2 shows the various screens of the application during the experiment.

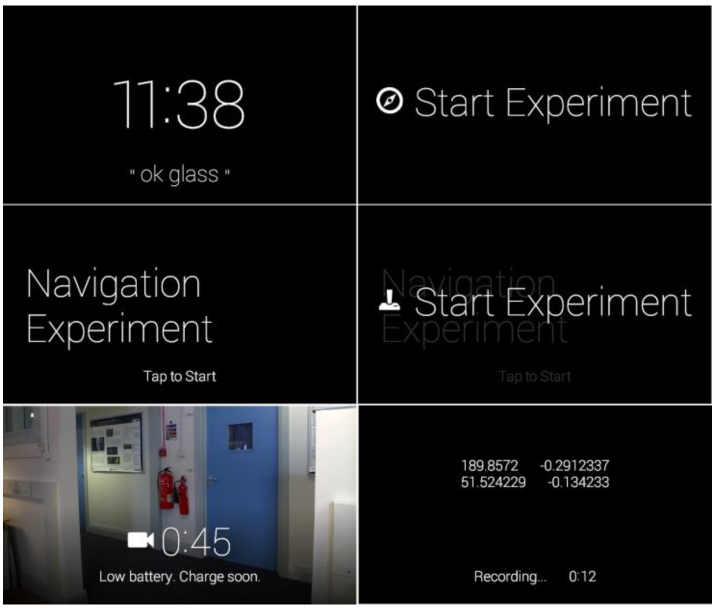

Figure 2 Screenshots of Google Glass application flow

3. Results and Discussion

The experiment collected two sets of data: 1) qualitative written instructions and; 2) quantitative gaze vector tracking.

Previous navigation studies using text analysis were constrained predominantly in virtual environments or pre-recorded video (Partala et al. 2010; Miller & Carlson, 2011). In this study, written instructions were produced after the participant were free to navigate a specified route. The assumption here was that the written instructions should reflect the most pertinent and salient environmental cues within a participant's cognitive map. The written instructions from all the participants were collated and used to create word clouds to provide an initial exploratory visual of the results. Common English words as well as directional terms (such as left, right, and across) were removed. This allowed the identification of landmarks and reference points used for the navigation instructions.

Figure 3 Word cloud summarising written instructions

The results showed that road names and landmarks feature heavily. This is consistent with existing literature on terms used in navigation instructions (May et al., 2003). It is key to note here the use of OS StreetView as a base for the route map may have led to a distorted high usage of road names and road signs. It can be concluded, however, where buildings are visually homogenous and there is a lack of obvious landmarks, street names become more important, especially for egocentric navigation strategies using personal directional instructions.

Further, the preceding and following description of the world "building" were extracted as a proxy to assess the intrinsic salient properties of urban landmarks (Table 1). Each term was categorised into two groups: visual and semantic. It can be seen that visual, rather than semantic, characteristics and cues such as colour, shape, texture and architectural style are most relevant when navigating at street level.

Table 1 Summary table of words preceding and following "building"

Description	Count	Description Type
red	7	Visual
glass	5	Visual
columned	3	Visual
large	3	Visual
brick	3	Visual
big	3	Visual
classical style	2	Visual
grey	2	Visual
stone	2	Visual
huge	2	Visual
UCL Hospital	2	Semantic
pointy	2	Visual
black	1	Visual
quaker's	1	Semantic
impressive	1	Visual
church-like	1	Visual
dark	1	Visual
single-storied	1	Visual
UCL	1	Semantic
cube	1	Visual
cream	1	Visual
Rubin	1	Semantic
University of London	1	Semantic
unusual	1	Visual
slatted	1	Visual
UCL Women's Health	1	Semantic
green	1	Visual
fancy	1	Visual
old	1	Visual
giant	1	Visual
grand	1	Visual
UCL Engineering	1	Semantic
small	1	Visual
white	1	Visual
Grant Museum of Zoology	1	Semantic

In this study, a gaze vector is defined as a geodesic line which most accurately represents the shortest distance between the participant location and the gaze position. The data collected from the participants were corrected for GPS error using individually digitised routes from the video. The median adjustment of 5.773m is consistent with the average median error of 8m for iPhone's integrated positioning technologies (Zandbergen, 2009). Each data point was also corrected to have a viewing height of 1.65m and the gaze vectors were then mapped using the orientation, pitch and locational data collected from the Google Glass. An assumption was made that the gaze position is calculated to be the first building or built structure from OS MasterMap[‡] that the gaze vector intersects within 200m (Figure 4). This was necessary as the device lacked true eye-tracking.

[‡] OS MasterMap Topography Layer Building Height Attribute was used to supplement this process.

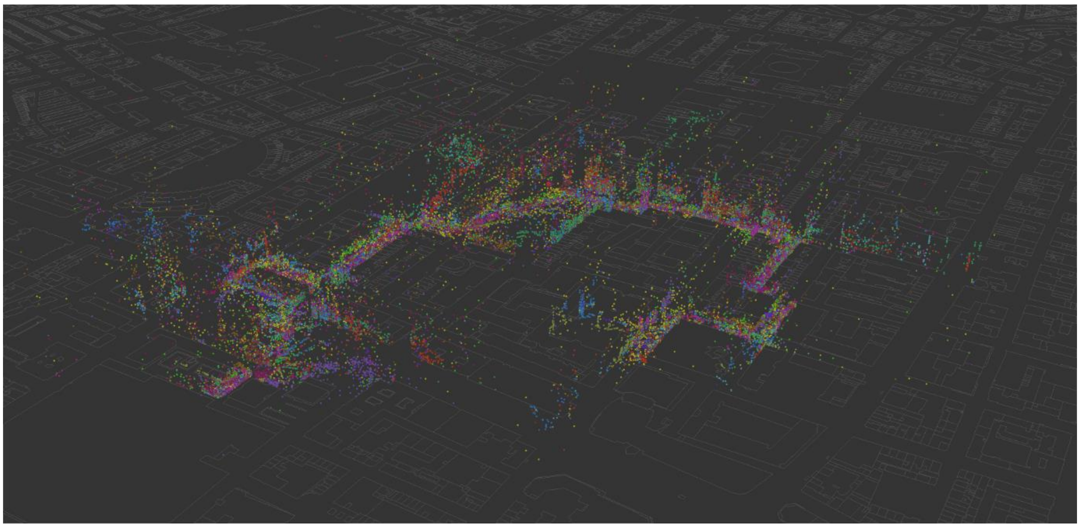

Figure 4 Gaze Positions of All Participants mapped in 3D

Initial exploratory statistics showed that gaze height was predominantly between 1.65 – 7.5m, with navigators focusing around eye-level to the first two stories. Roof characteristics, though useful for completeness and shape, are largely irrelevant pedestrian navigation as they could not be seen. The ideal subsequent analysis for this study would be to intersect gaze position points with a true 3D model, thereby identifying the exact features viewed while navigating every 0.125s i.e. the window on the ground floor or a street sign at an intersection. This, however, was not possible as a true 3D building city model with a high enough geometric and semantic detail was not available. This limitation stems not from the proposed methodology or data but rather the inherent deficit in suitable 3D GIS tools and datasets. An alternative exploratory analysis was carried out using 3D heat maps.

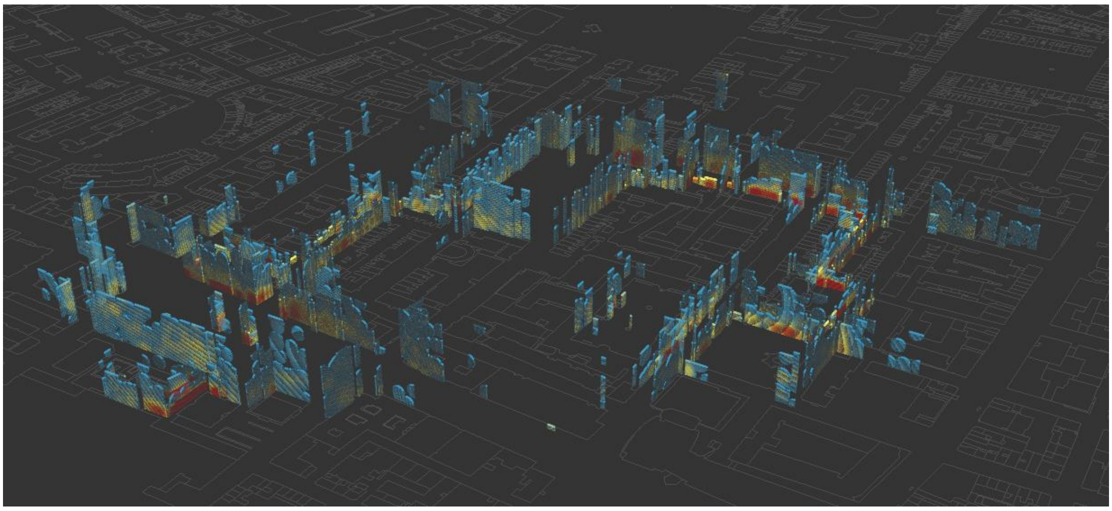

Figure 5 Gaze Positions Point Density Function of All Participants mapped in 3D

Every gaze positions was passed through a point density function to create a heat map for each building façade using the Gram-Schmidt orthogonalization process [see Wong (2012) for more detail] (Figure 5). A similar clustering approach was adopted in a web-based experiment identifying the features in a number of urban scenes which could be used in forming navigational instructions (Bartie et al., 2014). The 3D heat maps were overlaid against buildings in Google Earth as well as over textured buildings (Figures 6&7). The results showed that ground floor front-facing building facades were most examined followed by road signs. Buildings at decision points such as junctions were also used as landmarks.

Figure 6 Gaze Position Heat Map for Endsleigh Gardens in Google Earth

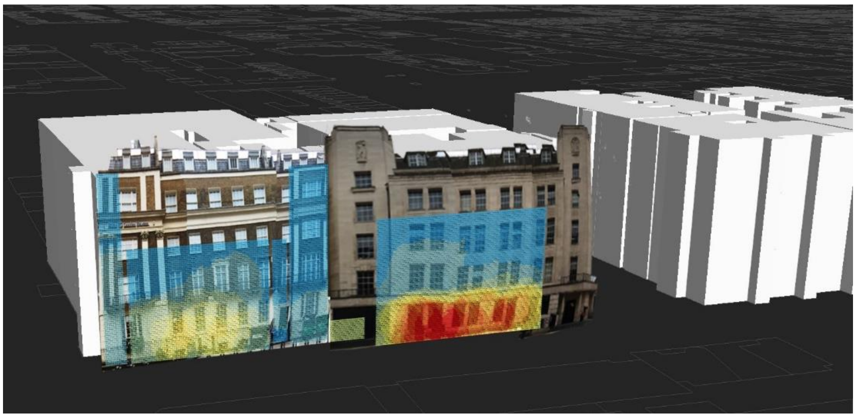

Figure 7 Gaze Position Heat Map for Endsleigh Gardens with building textures

The study and its subsequent analysis clearly demonstrates a severe shortcoming within the current state of GIS – where 2D is insufficient in analysing the data collected yet, true 3D data of adequate quality geometrically and semantically is unavailable. Further, there is a lack of GI systems able to perform geospatial queries in 3D which are readily available in 2D such as buffering, intersection and topological relations.

4. Conclusion

This study has provided a novel methodology in gathering requirements for a 3D navigation dataset. It outlines the first steps in an iterative study whereby the above recommendations would be implemented in developing a true 3D GIS dataset for navigation. Further work with participants from different age ranges and cultural backgrounds would be desirable to capture a representative sample of navigation strategies. In addition, true eye tracking and the availability of a 3D city model with full geometric and semantic attributes would enable the realisation of the full potential of the experiment. While the study demonstrates the possible value of 3D, it also shows the inherent deficiencies in the wider 3D field. In all, 3D is a multifaceted and ill-defined problem and it is unclear whether the benefits of the extra dimension outweighs its complexity. This study shows 3D is beneficial in the application of pedestrian navigation but argues existing technologies are incapable of delivering the envisaged true 3D navigation system. Where 2D maps works well on existing smartphone technologies, 3D navigation may require other enabling technologies such as a form of heads-up display or an ambient device. Regardless the direction 3D takes, what is key is that a user-centric design approach will ensure resulting outcome is effective, efficient, and enjoyable to use.

5. Acknowledgements

This project was funded and supported by the Ordnance Survey.

6. Biography

Kelvin Wong is an EngD research engineer at the UCL Centre for Virtual Environments, Interaction and Visualisation, Department of Computer Science, University College London. His research interests focuses on the challenges of deploying 3D geographic datasets at a national level with particular interests in usability, applications and data quality of 3D geographic information. Additional research relates to 3D visualisations and 3D requirements gathering.

Claire Ellul is a Lecturer in Geographical Information Science at University College London. Prior to starting her PhD, she spent 10 years as a GIS consultant in the UK and overseas, and now carried out research into the usability of 3D GIS and 3D GIS/BIM integration. She is the founder and current chair of the Association of Geographical Information's 3D Specialist Interest Group.

References

BARTIE, P., MACKANESS, W., PETRENZ, P. & DICKINSON, A. Clustering landmark image annotations based on tag location and content. Proceedings of RefNet Workshop on Psychological and Computational Models of Reference Comprehension and Production, 31st August 2014, 2014.

MAY, A. J., ROSS, T., BAYER, S. H. & TARKIAINEN, M. J. 2003. Pedestrian Navigation Aids: Information Requirements and Design Implications. *Personal Ubiquitous Comput.,* 7**,** 331-338.

MILLER, J. & CARLSON, L. 2011. Selecting landmarks in novel environments. *Psychonomic Bulletin & Review,* 18**,** 184-191.

PARTALA, T., NURMINEN, A., VAINIO, T., LAAKSONEN, J., LAINE, M. & VÄÄNÄNEN, JUKKA. Salience of visual cues in 3D city maps. Proceedings of the 24th BCS Interaction Specialist Group Conference, 2010. 428-432.

WONG, K. 2012. *3D Geographic Information and Solar Panel Positioning.* University College London.

ZANDBERGEN, P. A. 2009. Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning. *Transactions in GIS,* 13**,** 5-25.