

The Use of Eye Tracking in Search of Indoor Landmarks

P. Viaene¹, K. Ooms¹, P. Vansteenkiste², M. Lenoir², P. De Maeyer¹

¹Ghent University, Geography Department
Krijgslaan 281 (S8), 9000 Ghent, Belgium

²Ghent University, Department of Movement and Sport Sciences
Watersportlaan 2, 9000 Ghent, Belgium

{Pepijn.Viaene; Kristien.Ooms; Pieter.Vansteenkiste;
Matthieu.Lenoir Philippe.DeMaeyer}@UGent.be

Abstract. The detection of indoor landmarks remains a troublesome endeavour. The rise of more performant and user-friendly mobile eye tracking devices might offer a solution. A small-scale study was conducted in which a test population was given a navigational task and whereby eye movement measures and think aloud protocols were compared. The first results indicate that eye tracking has high potential for the specific task of identifying indoor landmarks, while thinking aloud offers minor additions to the information provided by eye tracking with respect to landmark identification.

Keywords: Think Aloud · Cognitive Processes · Wayfinding

1 Introduction

In line with the growing interest in indoor navigation and its challenges, indoor landmarks call for attention. These prominent elements in an environment enable an observer to locate himself and to set objectives like reaching a destination or selecting an optimal route [1]. Hence, indoor landmarks can serve as powerful wayfinding tools. Specifically, as part of view-action-pairs, they specify the location where a wayfinding action, which is needed to reach a certain destination, should take place [2]. In addition to this, landmarks are key elements in the construction of a spatial representation, which is central in our ability to navigate, as they anchor zones and form a hierarchical structure [3].

However, both in- and outdoors, it is not clear how landmarks should be detected and identified by researchers so that these objects can be studied and implemented in route instructions, maps and other wayfinding tools. A broad range of methods have been applied in the past with their specific (dis)advantages [4]. Some (e.g. [5, 6, 7]) tried to define landmarks by quantifying the features that contribute to the overall saliency of a landmark. However, these features and the way of quantifying the landmark's saliency vary. Moreover, the datasets on which these methods are based are in general not available for indoor environments.

With the development of more accurate and mobile eye trackers, measuring eye movements might be an adequate solution to identify indoor landmarks. First, the eye-mind hypothesis states that certain aspects of the gaze during a task may be analysed in order to examine cognitive processes [8]. While navigating, these processes are associated with the cognitive model of the environment, which is based on landmarks. The aspects that can be examined include the locus of the eye fixation and its duration. The locus indicates the element that is being processed internally even if subjects are not consciously aware of this and the duration is related, but not necessarily identical, to the time needed to encode and to operate on that element [8]. Second, landmarks are eye catching as they are highly distinguishable in their environment and differ from other objects based on visual, semantic and structural features [1].

2 Study Design

In order to assess the validity of the eye tracking method to detect indoor landmarks, the results of the eye movement analysis will be compared with the think aloud method, which is more commonly used to study cognitive processes related to (indoor) wayfinding (e.g. [9]) and therefore considered to be a valid representation of the cognitive processes related to the use of landmarks. A similar comparison was conducted by Spiers and Maguire [10]. However, they assessed to what extent the eye loci corroborated with the verbalizations in order to validate the verbal protocols. In this study, we wish to provide arguments for the validity of eye tracking itself.

Concurrent think aloud (CTA) is based on the analysis of verbal protocols formed by participants voicing their thoughts that come to mind while executing a problem-solving task [11]. In order to detect possible reactivity due to the extra workload of verbalizing, cued retrospective think aloud (CRTA) will also be part of this study. This method allows participants to execute the task silently – in this way not inducing an additional workload – and to verbalize their thoughts afterwards while watching a video recording of their performance on which their eye movements at the time are also displayed. These should cue the participants in revealing more about their thoughts at the time verbally [12]. However, as CRTA requires participants to remember information, it is possible that they forget important information [11, 12].

Twelve participants completed a route in a complex building¹ twice. The first time they had to follow the experimenter. The second time, they were asked to complete the same route independently. The experimenter only intervened if the participant was lost or asked for help. All participants wore the eye tracker during both completions of the route. Due to technical problems with the head mounted eye tracking device (iViewX HED by SMI), the recordings of three participants were excluded from the analysis. The remaining test population consisted of four subjects applying CTA during both traversals of the route and five applying CRTA based on the recording of the second traversal. The route itself had a total length of 440 meters and covered four floor levels. The participants, who had never been in the building, were made ac-

¹ University building: S8, Krijgslaan 281, 9000 Ghent, Belgium

quainted with thinking aloud before the experiment. Furthermore, they were told to verbalise everything, ranging from visual stimuli to feelings related to the navigational task and the building itself. Finally, the participants were aware that the goal of the study was related to indoor navigation and the use of landmarks.

The transcriptions of the verbal protocols were analysed with the aid of Elan EUDICO Linguistic Annotator (version 4.6.2). The protocols were split into verbalisation segments (e.g. one landmark referral, one explanation, one silence) and each segment was attributed with a time interval. The eye tracking data was analysed by using BeGaze 3.4. All fixations were transferred to a reference image that displayed 25 landmark categories (attributed with areas of interest) by using the semantic gaze mapping tool offered by SMI. Finally, verbalisations were compared with the eye movements (i.e. fixation locus and duration) around the same point in time.

3 Results

In total, 59 % (58 % (CTA), 61 % (CRTA)) of the verbalisation segments did not refer to a structural or object landmark. This accounted for 68 % (68 % (CTA), 69 % (CRTA)) of the observation time. A fourth of this 59 % consisted of segments containing additional information (e.g. explanations). This quarter was equal to 288 verbal utterances from which 118 were considered to be completely irrelevant for this study. This means that the information content of 170 relevant verbalisation segments, which represented 16 % of all segments, was not part of the eye tracking data. Following, the remaining three quarters of the 59 % represented silences and corresponded to 57 % (66 % (CTA), 56 % (CRTA)) of the total observation time.

With respect to the 41 % of segments that did refer to a potential landmark, the following can be said. On average, 69 % (71 % (CTA), 66 % (CRTA)) of the mentioned potential structural and object landmarks were clearly fixated on. On the other hand, 13 % (12 % (CTA), 14 % (CRTA)) of the described landmarks were not visible at the moment they were verbalised. The remaining 18 % (16 % (CTA), 20 % (CRTA)) represented the number of indoor landmarks that could not be unambiguously identified solely based on the eye tracking data and therefore verbalisations were needed for a true determination.

We now turn to the locations where landmarks were most needed, namely locations where a change of direction took place and where multiple directional possibilities were present. The most fixated on landmark categories at these decision points are shown in Table 1. Often a single object caused the rise in fixations for a specific category. These object landmarks, defined as elements that are independent of the building's structure [3], are listed in Table 1 as well. As it is not clear how people visually perceive structural elements (i.e. staircases and corridors), these elements were excluded from the eye tracking analysis. The fixated on objects at the seventeen decision points were compared with the objects mentioned at these locations in the thirteen verbal protocols. In 59 cases there was a match, while other objects were mentioned 73 times. Often, these other objects were staircases (34 times). Finally, 89 times there were no referrals to objects.

Table 1. The most fixated on landmark categories and related object landmarks based on the maximum average fixation count, maximum average fixation time and fixation time maximum.

DP	landmark category	object landmark
1	door (route)	<i>grey double door</i>
2	other / route indicator	<i>exhibition display</i>
3	route indicator	<i>sign ("Geography")</i>
4	door (route)	<i>brown double door</i>
5	window	<i>window and view</i>
6	door (route) / other	<i>pair of sticks / car batteries</i>
7	door (route)	<i>brown doors with windows</i>
8	ornament	<i>big plant</i>
9	elevator	<i>red elevator</i>
10	poster	<i>wooden information board</i>
11	door (other)	<i>grey double door</i>
12	door (other)	<i>glass main entrance</i>
13	route indicator / other	<i>sign ("Paleontology")</i>
14	door (other)	<i>brown double door</i>
15	window / route indicator	<i>window and view</i>
16	door (route)	<i>brown double door</i>
17	door (route) / poster	<i>single door</i>

4 Discussion and Conclusion

The general comparison between eye recordings and verbal protocols leads to two findings. First, a considerable share of the information originating from the think aloud method is not deductible by tracking eye movements, namely a quarter of all verbalisations: 13 % non-visible landmarks and 16 % relevant verbalisations without referral to landmarks. However, the latter is not considered to be a loss of information since these do not contain references to potential indoor landmarks, given that the goal of this study is to determine if eye tracking could be used specifically to identify indoor landmarks. Second, although all participants stated that they did not experience difficulties with respect to voicing their thoughts, the think aloud method did not supply information during more than half of the observation time. Pointing out that the quality of verbal protocols depends on the skills of the respondent. Respondents sometimes only verbalize part of their thoughts or have difficulties translating their thoughts into words [11]. Moreover, subjects can only provide data on processes that they are aware of [10]. In contrast, eye tracking provided data continuously.

With respect to the most fixated on objects, there is a poor resemblance. However, when neglecting referrals to staircases, as fixations on staircases were not seen reliable, one can conclude that there were only 39 mismatches. Consequently, there were no referrals to objects in 123 of the cases, which is in line with the observation that thinking aloud does not supply information in more than half of the observations. Furthermore, the fact that staircases were often mentioned does not automatically

mean that these structural elements were remembered as wayfinding aids. An explanation might be found in the physically perceivable interaction with these elements [3]. Finally, there were no indications that CTA caused reactivity that had significant effects on task performance or concentration.

In conclusion, the results indicate that eye tracking can provide qualitative and complete data which can be used to identify indoor landmarks. Although eye tracking captures most information relevant for the identification of landmarks, it is advisable to record verbal protocols which can be consulted to clarify specific fixations in order to obtain a more complete outline of potential landmarks. However, having the time-consuming analysis of verbal protocols in mind, these should not be the subject of a separate secondary analysis since the added value is limited.

References

1. Sorrows, M., Hirtle, S.: The nature of landmarks for real and electronic spaces. In: Freska, C. and Mark, D.M. (eds.) *Spatial information theory. Cognitive and Computational Foundations of GIS*. pp. 37–50. Springer-Verlag, Berlin, Germany (1999).
2. Lovelace, K.L., Hegarty, M., Montello, D.R.: Elements of Good Route Directions in Familiar and Unfamiliar Environments. *Spatial information theory. Cognitive and Computational Foundations of GIS*. pp. 65–82. Springer-Verlag, Berlin, Germany (1999).
3. Stankiewicz, B.J., Kalia, A.A.: Acquisition of structural versus object landmark knowledge. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 378–390 (2007).
4. Sefelin, R., Bechinie, M., Müller, R., Seibert-Giller, V., Messner, P., Tscheligi, M.: Landmarks: yes; but which?: five methods to select optimal landmarks for a landmark-and speech-based guiding system. 7th international conference on Human computer interaction with mobile devices and services. pp. 287–290. ACM Press, Salzburg, Austria (2005).
5. Raubal, M., Winter, S.: Enriching Wayfinding Instructions with Local Landmarks. In: Egenhofer, M.J. and Mark, D.M. (eds.) *Geographic Information Science. GIScience 2002*. pp. 243–259. Springer-Verlag, Berlin, Germany (2002).
6. Fang, Z., Li, Q., Zhang, X., Shaw, S.: A GIS data model for landmark-based pedestrian navigation. *Int. J. Geogr. Inf. Sci.* 26, 1–22 (2011).
7. Elias, B.: Determination of Landmarks and Reliability Criteria for Landmarks. Fifth workshop on progress in Automated Map Generalization Paris. pp. 1–12. , Paris (2003).
8. Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. *Cogn. Psychol.* 8, 441–480 (1976).
9. Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., Knauff, M.: Up the down staircase: Wayfinding strategies in multi-level buildings. *J. Environ. Psychol.* 26, 284–299 (2006).
10. Spiers, H.J., Maguire, E. a: The dynamic nature of cognition during wayfinding. *J. Environ. Psychol.* 28, 232–249 (2008).
11. Van Elzakker, C.P.J.M.: *The Use Of Maps In The Exploration Of Geographic Data*. Koninklijk Nederlands aardrijkskundig genootschap, Faculteit geowetenschappen, Universiteit Utrecht / Internationaal Instituut voor Geo-Information Science and Earth observation, Utrecht/Enschede (2004).
12. Van Gog, T., Paas, F., van Merriënboer, J.J.G., Witte, P.: Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *J. Exp. Psychol. Appl.* 11, 237–244 (2005)