I Can Tell by Your Eyes! Continuous Gaze-Based Turn-Activity Prediction Reveals Spatial Familiarity

Negar Alinaghi
Geoinformation, TU Wien, Austria

Markus Kattenbeck
Geoinformation, TU Wien, Austria

Ioannis Giannopoulos
Geoinformation, TU Wien, Austria
Institute of Advanced Research in Artificial Intelligence (IARAI), Wien, Austria

Abstract

Spatial familiarity plays an essential role in the wayfinding decision-making process. Recent findings in wayfinding activity recognition domain suggest that wayfinders' turning behavior at junctions is strongly influenced by their spatial familiarity. By continuously monitoring wayfinders' turning behavior as reflected in their eye movements during the decision-making period (i.e., immediately after an instruction is received until reaching the corresponding junction for which the instruction was given), we provide evidence that familiar and unfamiliar wayfinders can be distinguished. By applying a pre-trained XGBoost turning activity classifier on gaze data collected in a real-world wayfinding task with 33 participants, our results suggest that familiar and unfamiliar wayfinders show different onset and intensity of turning behavior. These variations are not only present between the two classes –familiar vs. unfamiliar– but also within each class. The differences in turning-behavior within each class may stem from multiple sources, including different levels of familiarity with the environment.

2012 ACM Subject Classification Computing methodologies → Activity recognition and understanding; Computing methodologies → Supervised learning by classification

Keywords and phrases Spatial Familiarity, Gaze-based Activity Recognition, Wayfinding, Machine Learning

Digital Object Identifier 10.4230/LIPIcs.COSIT.2022.2

Supplementary Material *Dataset (train and test)*: https://doi.org/10.48436/f0chy-11p06

Acknowledgements We would like to thank Antonia Golab (Vienna University of Technology) for collecting the valuable data used for this work.

1 Introduction

Wayfinding as a general concept is outlined as "the most prominent real-world application of spatial cognition" [40]. It is an ongoing decision-making process consisting of several tasks, each of which requires cognitive resources [3]. The cognitive demand for each task heavily relies on environmental and user-related features. Theoretical reasoning (e.g. [40, 14, 15]) and empirical evidence (e.g. [31, 23, 25, 12] suggest that *familiarity with the environment* as a particular state of spatial cognition plays an important role in the wayfinding decision-making process. Our previous research on the prediction of wayfinders' turning activity, as an actual realization of a series of spatial decisions, introduced *familiarity* as the most important/ influential feature for the prediction model [2].

OSM Open Street Map
ML Machine Learning
POI Point of Interest

15th International Conference on Spatial Information Theory (COSIT 2022).
Editors: Toru Ishikawa, Sara Irina Fabrikant, and Stephan Winter; Article No. 2; pp. 2:1–2:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Motivated by our previous finding [2], presenting *familiarity with the environment* as the most important feature in turning activity prediction, we decided to more thoroughly look into this feature during the decision-making process for the turning activity. After receiving any instruction, wayfinders try to match the conveyed spatial information to the physical environment and decide for their turning action (i.e., whether to turn left/right or continue straight ahead). Trying to understand the undergoing cognitive processes in this *matching-to-action* phase may reveal aspects of spatial cognition, including spatial familiarity. Understanding the behavioral correlates of familiarity in this phase of decision-making is a valuable theoretical contribution in the spatial familiarity domain and provides a fertile ground for cognitively engineered spatial systems, e.g., more context-aware navigation aid systems.

In this paper we take a step into this direction: We demonstrate that familiar and unfamiliar wayfinders, exhibit differing behaviors while reacting on navigational instructions to decide whether to turn or continue straight (this we call "turning-activity"). On this account, we monitor the wayfinders' gaze behavior in relation to their turning-activity, i.e., during the matching-to-action phase of the decision-making process, using a mobile eye tracker and a high-precision GNSS receiver. We report on an in-situ pedestrian wayfinding study ($N = 33$), during which participants walked two routes: one located in an area with which they were familiar, whereas another route located in an area they were unfamiliar with. The familiarity with the environment was reported by participants in a multi-step procedure online study prior to the in-situ experiment (see Section 3.1 for more details). We adapted the pre-trained XGBoost model taken from our previous research [2] and applied it to the gaze behavior of these two groups of wayfinders, within the matching-to-action phase.

Our analysis provide evidence that detecting gaze-behavioral differences concerning the turn activity reveals the familiarity of the wayfinders as a binary measure. In addition to that, as the distribution of probabilities predicting turn vs. no turn activities varies within both groups, familiar and unfamiliar participant, alike, our findings may lead to a potential proxy for the estimated level of familiarity.

2 Related Work

Given the aforementioned research goals, we first review previous findings on familiarity aspects of wayfinding and continue with wayfinding studies that examine gaze behavior during decision-making. Finally, we report on the prior work on using Machine Learning (ML) methods for activity classification based on eye-tracking data.

2.1 Familiarity Aspects in Wayfinding and Beyond

"We reason that a navigator's search behavior and search strategy will be heavily influenced by their degree of familiarity with the environment." [40]. Investigating the effect of familiarity on wayfinding and environmental learning, as well as understanding the variables that contribute to the development of a sense of familiarity, constitute the majority of theoretical research in the domain of spatial familiarity [1]. Up until now, however, we lack a comprehensive conceptual definition for "*spatial familiarity*" [15]. As a consequence, assessing and modeling familiarity continue to be open research problems. The majority of attempts taking familiarity into consideration rely on self-reported measures, e.g., customized questionnaires and sketched maps (see e.g. [6, 19]) and these studies most often use a binary definition of environmental familiarity (see e.g., [43, 25, 2]). Approaches exploiting behavioral data benefit from various sources ranging from mobile phone GPS tracks (see.

e.g., [37]) to social media (e.g., [38]) and recently also gaze-behavior ([25]). Having said this, empirical studies addressing environmental familiarity have been conducted both indoors [18], outdoors [25]) and in virtual environments [12], alike.

Only recently, has the classification of binary familiarity based on behavioral correlates gained momentum. For instance, Gokl and colleagues [12], reported 51.87% to 65.70% accuracy for a binary classification on familiarity performed on behavioral data collected from an avatar-based VR study. Savage and colleagues in [34], used place-visits in Foursqure application and combined it with user-contextual information taken from Facebook profiles and GPS trajectories to score the level of familiarity with places using Bayesian techniques. Liao and colleagues in a very recent paper [25], performed a Random Forest binary classification (data was initially collected on a Likert-like scale) for familiarity using the gaze-behavior collected in a real-world navigation task with 38 participants and reported 70% to 81% accuracy.

2.2 Gaze-based Activity Recognition

As one of the most promising ways of gaining access to people's cognitive state, gaze, has received a great deal of attention in the GIScience research domain (see [21, p. 2–9] for an overview): Its utility has been explored in many applications, for instance, tailoring wayfinding assistance systems to individual's demands (see e.g. [11]), and numerous wayfinding studies conducted in real-world or virtual reality environments ([35], [8] and [39]).

Having said this, the importance of gaze in task prediction has seen great interest for decades, starting with the seminal works of Yarbus [41] and Buswell [5]. For instance, Kiefer and colleagues [20] performed gaze analysis for automatic recognition of map users' activities by introducing novel gaze features based on string sequence analysis, and reported 78% accuracy in classifying six common map activities (free exploration, global and focused search, route planning, line following, and polygon comparison). Bulling and colleagues [4], also applied a Support Vector Machine (SVM) on gaze features for recognizing five office-related activities (e.g. copying a text, reading a printed paper etc.) in an eight-participant study and reported an average precision of 76.1%. In addition to task prediction, researchers have tried to predict the underlying cognitive or even demographic aspects which are reflected in various tasks. For instance, Henderson and colleagues [17], tried to predict the cognitive workload of 12 participants performing various tasks e.g., scene memorization, text reading etc., using a multivariate pattern classification on the recorded eye movements with 68% to 89% accuracy. Galdi and colleagues [9], studied the applicability of gaze analysis for gender and age categorization by applying Adaboost and SVMs on data collected from 112 different observers.

Similar recent research in our domain includes the work of Alinaghi and colleagues [2]. The authors trained an XGBoost classifier on gaze data acquired from a real outdoor wayfinding study. They reported the best performing model (with 91% of accuracy for predicting the turning activity of wayfinders with three classes of No-Turn, Turn-Left, and Turn-Right) trained with the gaze data from the last three seconds before the turning action was performed. Liao and colleagues [24] reported an accuracy of 67% using a Random Forest classifier on 38 recorded eye movements during navigation, to predict five tasks, namely self-localization, environment target search, map target search, route memorization, and walking to a destination.

2.3 Machine Learning Models for Activity Recognition

The related work reviewed so far, already indicates the prevalence of SVMs and tree-based ML techniques in performing classification tasks based on eye movement data. For instance, Bulling and colleagues [4], and Shiga and colleagues [36] both used SVMs for classification of general office tasks and everyday tasks based on gaze data and reported reasonable results. Liao and colleagues reported promising results by applying Random Forest both for task recognition ([24]), and familiarity prediction ([25]). Liu and colleagues' results from Random Forest in [26], where they examined differences in gaze behavior related to the regularity of road patterns and signage, are also encouraging.

Although SVMs and tree-based techniques provide promising results in gaze-analysis, recent studies report higher accuracies achieved by ensemble learning, e.g., Gradient Boosted Trees, across different domains. For instance, in their gaze-based turning activity prediction, Alinaghi and colleagues [2], compared SVM-RBF, CART, and Random Forest in a pilot testing phase and reported that XGBoost achieved higher accuracy. Similarly, Zhang and colleagues [42] provide evidence that XGBoost-based indoor activity recognition algorithm outperforms both other ensemble learning classifiers and single classifiers in indoor activity classification (achieving 84.41% accuracy). In a different domain, Mathur and colleagues [27] found XGBoost superior to eight ML models as well as two deep learning models (long short-term memory networks and temporal convolutional networks), with an accuracy of 69% and an AUC of 72% when detecting users' empathy while listening to a narrative robot.

3 Detecting Familiarity Based on Turn Activity

With a similar approach as [17, 9], where trained models for specific task recognition problems were deployed to predict user-related characteristics, we have also used a pre-trained turning-activity classifier for familiarity detection. This section provides information on our analysis for familiarity detection from turning activity behavior. We start with a short description of the human-subject in-situ study as we use the same dataset as [2] which was collected to address several research questions. We, then, move on to explain our feature extraction approach. Finally, we describe the prediction model that we deployed for our familiarity detection and provide details on how we came upon the familiarity patterns by monitoring the turning activity behavior during the matching-to-action phase.

3.1 Data Collection

The data used in this paper is part of a larger data collection effort in 2020 addressing human spatial behavior in real-world wayfinding scenarios and is first presented in [13][1]. In a within-subject design study, participants were required to walk two routes each, one of which was located in an area they were familiar with, while they were unfamiliar with the other. Environmental familiarity was collected during an online study: Participants were asked to indicate areas in which they would be able to find their way easily without any kind of wayfinding assistance This was used as proxy for being familiar. Subsequently, participants were asked to pinpoint familiar places therein. Two of these places were randomly chosen based on the condition of being between 0.9 and 1.3 km apart in order to ensure a reasonable duration of the experiment, and participants provided their preferred route between these

[1] Parts of the data used in the current paper, will be made available at: `https://geoinfo.geo.tuwien.ac.at/resources/` (DOI: `10.48436/f0chy-11p06`).

places. All participants walked the familiar route they provided and were randomly assigned an unfamiliar route which was located in a polygon with no overlaps with the polygons a participant indicated as familiar.

For either route, participants had to find their way by means of auditory, German-language, landmark-based[2], turn-by-turn route instructions, which were provided to them on demand and as many times as they requested. In order to ensure the demand of route instructions when walking a familiar route, participants were instructed that they might be guided on a route different to the one they actually provided. Participants' behavior during the experiment was monitored with a mobile eye-tracking device (PupilLabs Invisible; 200Hz recording frequency), high precision GNSS receiver (PPM 10-xx38) that tracked their position in time and a head-mounted Inertial Measurement Unit (IMU), whose data we do not use in this study. Additionally, participants were given a small clicker-device, to request an instruction by simply pressing a button. In total, $N = 52$ people (27 female and 25 males, $M(age) = 26$ years, $SD(age) = 8.3$) participated in the outdoor experiments resulting in $N = 104$ trials out of which $N = 32$ were retained for further processing in this paper (18 trials: equipment malfunction or participants not cooperating; 54 trials: not suitable given our research question, see Section 3.3 for details).

3.2 Data Preparation

To obtain the gaze-related data at each decision-point, we first matched each of the Open Street Map (OSM)-driven junctions along a route to their corresponding GPS point by drawing a ray oriented along the segments of the given intersection and intersecting it with the GPS trace. As the result of this step, we obtained the corresponding timestamps at which each junction was reached.

Subsequently, we approximated the *Start* points (i.e. *when an instruction was given for the first time*) and *End* points (i.e. *when a junction, for which the instruction was given, was perceived as junction*) of the matching-to-action phase in our analyses. From now on, we refer to each pair of these points as *test-sample*. Figure 1 A, visualizes these points. Start points were found by synchronizing the clicker-output with GPS timestamps. This point is depicted as a green speaker icon in Figure 1 A. The yellow icons in this figure indicate the points in time when the same instruction was requested for a second time (again, based on the synced clicker and GPS timestamps). Since the perception of decision *points* in reality do not exactly equate to the junction point (e.g., we do not decide to turn left to a street right in the middle of a crossing), we model the decision *situations* by using OSM building footprints to find all intersection boundaries[3]. These boundaries are presented in Figure 1 A, as hatched polygons located at each junction (orange: *No-Turn*; blue: *Turn*). From now on, we will refer to the intersection point of the GPS trace with this boundary as *decision-point*. Figure 1 A represents these decision-points with green and blue circles denoting the start and end of each segment in the test-sample.

[2] Points of Interest (POIs) were used as landmarks and chosen according to the algorithm described in [33]. For a detailed description on how these were double-checked, see [13]. Route instruction pattern: TURN LEFT [IMPERATIVE] AT CAFE RITTER. [LANDMARK AND, HENCE, LOCATION OF TURNING ACTION].

[3] If no buildings were located at a junction, we found the boundary of the intersection by using a threshold of $3.75secs$ from the projected junction point (i.e. $\approx 5m$, based on $4.5km/h$ avg. walking speed [22].

3.3 Sliding-Window Feature Extraction

As the highest accuracy for the XGBoost turn-activity classifier in our previous research [2] was reported when trained three seconds before the actual turn occurred, we selected three seconds as our window size. The step distance was also set to three seconds, avoiding any overlaps between windows. This is based on findings suggesting that non-overlapping windows deliver comparable recognition accuracy while majorly reducing the required training computations and memory usage [7].

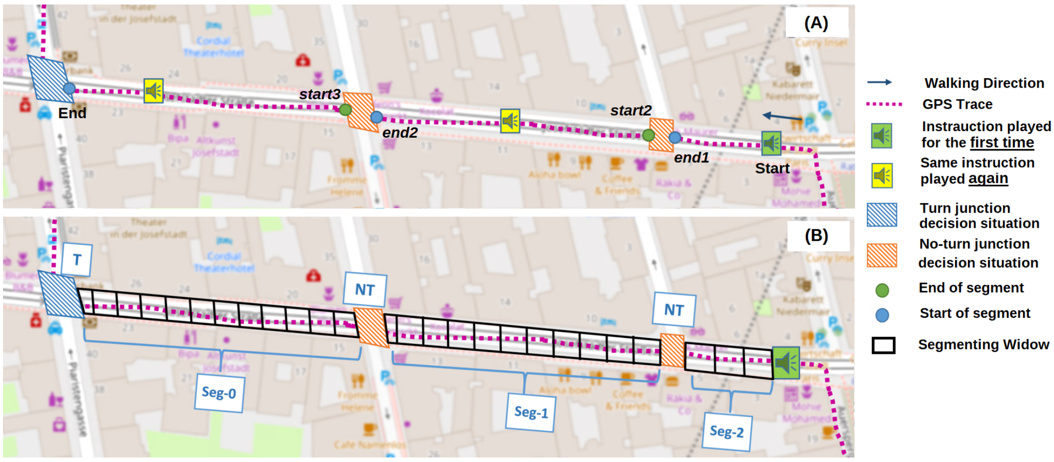

Figure 1 Figure **A**, representing one *test-sample*, visually defines the *Start* and *End* points of the test-sample as well as the middle *start(i)* and *end(i)* points of the segments. As schematically presented in **B**, each test-sample is then segmented with a sliding window approach into non-overlapping windows of three seconds duration. Since we want to analyze the results segment-wise, we group the windows belonging to one segment by segment id as *Seg-i*.

With the method explained in 3.2, we extracted all the *Start* and *End* points per turning junctions in all trials, with respect to two main conditions. First, as the experimental design allowed participants to ask for instructions as often as they needed, we decided to exclude all the repetitions and only keep the first point in time when the instruction was given. Second, test-samples should at least have one *No-Turn* junction (denoted in Figure 1 B as *NT*). This decision is made to have enough analysis time and more decision situations in the matching-to-action phase. This ensures that we always have at least two segments in each test-sample. The naming convention for identifying these segments is as follows: *Seg-**n*** with n starting from zero always indicating the final segment right before the *Turn* junction. This setting, i.e., starting each test-sample always from the initial instruction point and excluding the data before that, ensures a clean test-sample and is inline with the experimental design in which both un/familiar groups did not know the route in advance and had to relay on the given turn-by-turn instructions to reach the destination. Figure 1 B, represents one of these clean test-samples for which two *No-Turn* (*NT*) junctions were passed before the *Turn* (*T*) junction was reached. This results in three segments with ids starting from *Seg-0*, to *Seg-2* which is the first segment right after the instruction was given. After selecting the test-samples according to these conditions, we ran a non-overlapping sliding window (with $window - size = 3sec$ and $step - size = 3sec$) approach for segmenting the eye-tracking data. Figure 1 B, visually present this approach.

3.4 Detecting Familiarity by XGBoost Turning Activity Classifier

The trained XGBoost turning-activity prediction model from our previous research [2] is modified and used for familiarity detection purpose in this paper. XGBoost is one of the Gradient Boosted Tree algorithm implementations, which allows parallel tree boosting for unfolding very complex patterns in a highly efficient and scalable manner.

The model was trained on 31 features including 28 fixation- and saccade-based gaze feature (i.e., fixation frequency, min/max/mean/variance of fixation- duration/dispersion/dispersionX and dispersionY; and saccade frequency, min/max/mean/variance of saccade-amplitude/duration, skewness of saccade amplitude, and g-l ratio which is the ratio between long and short saccades), 2 environmental features (i.e., number and skweness of street segments at each junction) and 1 user-related feature (i.e., familiarity with the environment as a binary measure) for 1335 junctions acquired from the same dataset used in [2] (see Section 3.1). In that paper [2], the highest accuracy of turning-activity prediction (91.4%) was achieved when the model was trained with the data from the last three seconds before the turning action. By analyzing the SHAP values, we came upon *familiarity* as the most important feature for the model.

As explained in Subsection 3.3, we segmented the data within each test-sample into three seconds windows to use this model. To investigate the effect of familiarity on turning activity, we customized this pre-trained model in two ways: We modeled the turning activity as a binary measure: *Turn* vs. *No-Turn*, i.e. we did not distinguish between left and right turns. In addition to that, we turned the model into a *probabilistic classifier* in order to gain the probability distribution over the two classes. It has been suggested in statistics that such *posterior probabilities* are required “when a classifier is making a small part of an overall decision or the outputs need to be combined for the overall decision” [32]. Our following familiarity analysis also requires the investigation of these small parts of an overall classification decision.

Therefore, the design decisions presented here are based on our initial assumption that the expected variations in the posterior probabilities can be considered as a proxy of spatial familiarity regarding the matching-to-reaction phase of the turning activity. In order to visually inspect patterns for *familiar* and *unfamiliar* cases, we plotted these probabilities in the matching-to-action phase.

4 Results

This section provides the results of our familiarity detection based on turning activity behavior during the matching-to-action phase. The two subsections outline the results from three different angles: the primary outcome of our analysis representing the overall difference between *familiar* and *unfamiliar* wayfinders with regards to their turn-activity behavior is presented in subsection 4.1; subsection 4.2, however, sheds light on the results of a case-wise between class comparison (familiar vs unfamiliar) and the result of a case-wise within class comparison, alike.

4.1 Overall Familiarity Detection

Figure 2 represents the median percentage of windows in which a turn is predicted, aggregated along each segment across all trials. To analyze the computed probabilities across all cases, we aggregated the results by computing the median percentage of windows belonging to *Turn* class in each segment. This percentage represents the number of windows for which the probability of class *Turn* was higher than the probability of class *No-Turn*. Since the segments (i.e., street segments belonging to different routes) within each test-sample have different lengths, we computed the percentages *segment-wise*. This ensures that the final result is normalized and comparable on one scale. When interpreting the figure, it is important to keep in mind that test-samples vary concerning the number of segments they have: For instance, depending on the route and the first time someone asked for instruction, one test-sample ended up having three segments while another included four. However, the second condition we had set for selecting test-samples assured the presence of at least two segments (i.e., *Seg-0* and *Seg-1*): As explained above, we had chosen the test-samples in a way such that at least one *No-Turn* junction was passed before reaching the *Turn* junction for which the turn-instruction was given. In Figure 2, the notation n represents the number of trials for which we have calculated the median percentage.

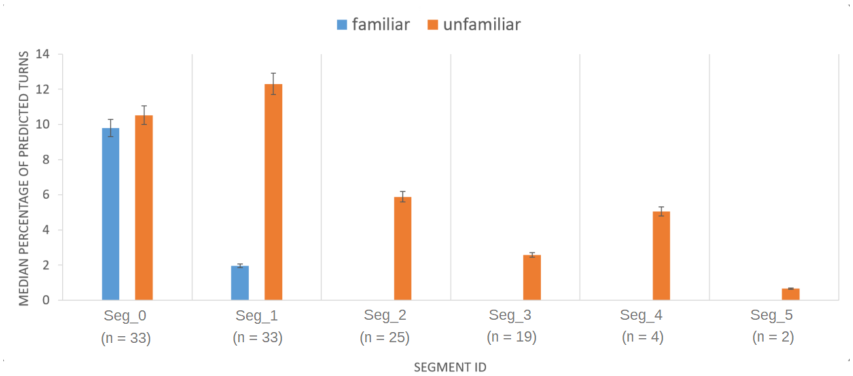

Figure 2 This plot represents the aggregated results, measured by the median percentage of windows per segment for which the class *turn* is predicted, for the two targeted groups: *Familiar* vs. *Unfamiliar* wayfinders. Note that not all the test-samples have all the segments and as a result the same number of time windows (due to the difference in the number of junctions and instruction-click points per route). The number of test-samples per segment is denoted by the letter n and error bars indicate the 95% confidence limit. The results suggest that familiar wayfinders are more confident in matching the turn instruction to the spatial environment.

4.2 Case-wise Comparisons: Between- and Within-Class

Figure 3 illustrates the resulting plots for a test-sample taken from one route walked by a familiar and an unfamiliar wayfinder. Since familiar and unfamiliar routes traveled by wayfinders differ in terms of the street network, length, urban complexity, etc., we compare turning behavior on a single route walked by two different wayfinders. The *x-axis* in Figure 3 represents multiple information for clearer understanding: the window-id, constellation of instruction-point (depicted in green) and junctions (distinguishing *Turn* and *No-Turn* junctions with blue and orange), and the segments outlined in purple for the last segment immediately before the turning junction and yellow for other segments. To visualize the turn behavior of each wayfinder, the probability of classes *Turn* (orange) and *No-Turn* (blue) within the matching-to-action phase are given for each window. The difference between familiar and unfamiliar wayfinders is apparent in this plot: The onset and the intensity of the variations in probabilities for the two plots are within different segments before the actual turn.

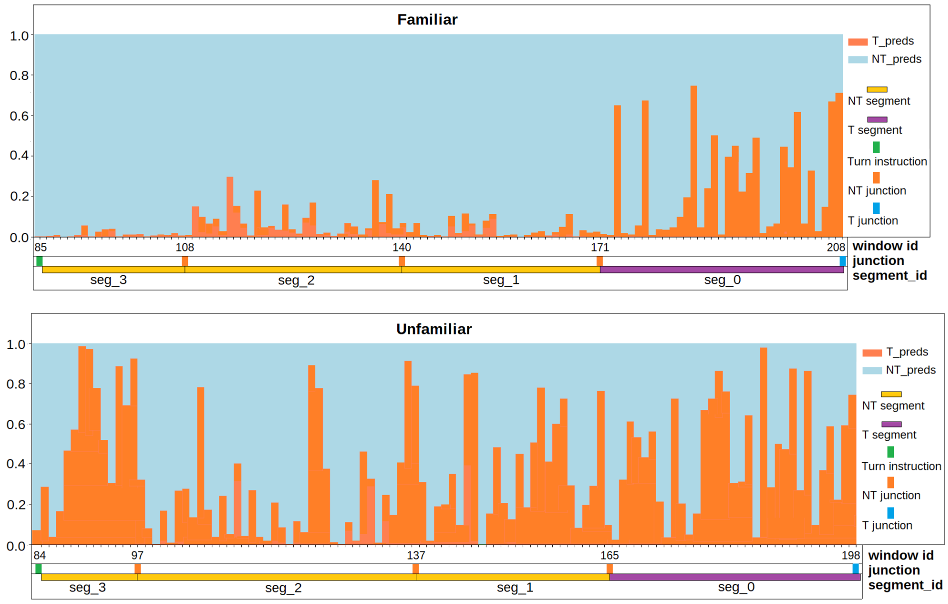

Figure 3 This figure contrasts the probability plots of one familiar and one unfamiliar wayfinder for a single route. The plots indicate that the variation in turning activity probability (i.e., turning-behavior) starts at different stages between the two classes: The familiar wayfinder demonstrates more variations only in the last segment before the turn-junction (*Seg-0*), while the unfamiliar wayfinder exhibit this behavior closely after receiving the instruction (*Seg-3*). Along the x-axis, a schematic view of the test-sample in terms of the constellation of instruction point, *NT* as well as *T* junctions is shown. Note that although the two diagrams represent the same route, they cannot be fully aligned due to different instruction-click times and wayfinders' walking speeds.

By investigating the probability plots per trial within each class (familiar and unfamiliar), we observed different distribution of turning-activity probabilities. The onset and intensity of turning behavior are equally different within familiar and unfamiliar wayfinders. As the in-situ experiment was designed so that different un/familiar wayfinders walked different routes, disentangling the impact of the route and individual differences might not be feasible. Coincidentally, however, one route was walked by two different unfamiliar wayfinders. This gave us a single case where the environment is fixed, and we could relate the patterns of turning behavior to individual differences. Figure 4, illustrates the probability plots for this case. The plots show differences concerning the distribution of the turning activity probabilities even within one familiarity class. A similar pattern is also present in both familiar and unfamiliar classes when different routes are considered. This is particularly remarkable for familiar cases: While the vast majority of cases showed a variation of probabilities only on the final segment, there were also cases showing these variations sooner, e.g., in *Seg-1* or *Seg-2*.

5 Discussion

We analyzed turning activity behavior in the *matching-to-reaction* phase, which starts right after an instruction is given until the corresponding junction is reached. We interpret the variations in turn probabilities as an indicator of *turning-behavior* stemming from the given instruction. We provide evidence that familiar wayfinders show a different behavior during

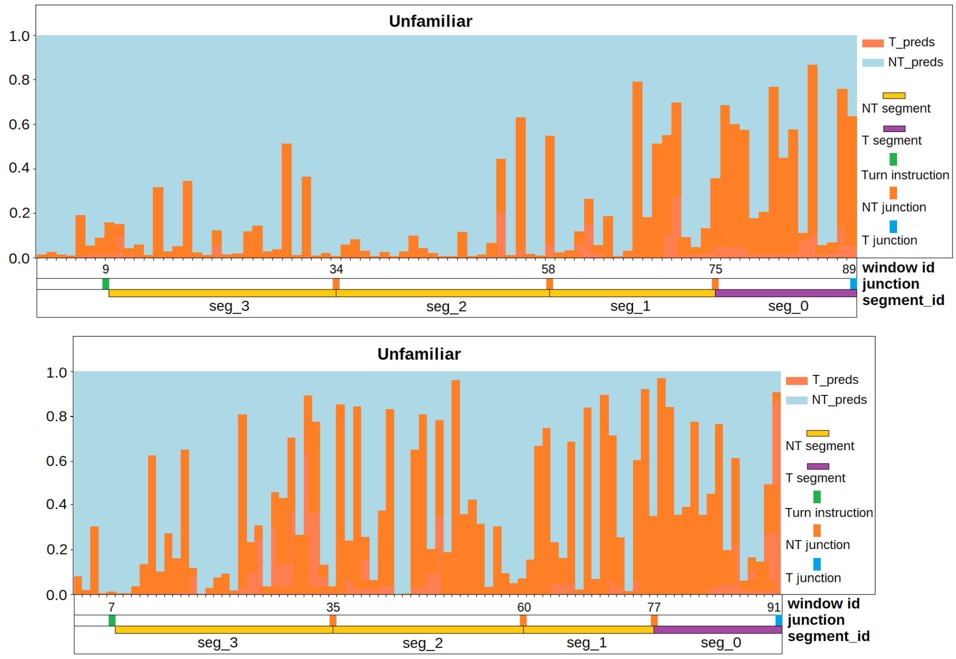

Figure 4 This figure shows the probability plots of an exceptional case in which two unfamiliar wayfinders coincidentally walked a single route. The plot indicates that the variation in turning activity probability (i.e., turning-behavior) starts at different stages. Along the x-axis, a schematic view of the test-sample in terms of the constellation of instruction point, *NT* as well as *T* junctions is shown.

the matching-to-reaction phase (high median percentage of predicted turns only in the final segment before the turning junction) of the decision-making compared to unfamiliar wayfinders (higher rate of turning behavior compared to the familiar group in almost all segments once an instruction was received).

While familiar and unfamiliar wayfinders can be distinguished reasonably well based on their gaze behavior before turns, our data suggest considerable within-class diversity. This was revealed when the results within each class were analyzed case-wise. Within both familiar and unfamiliar groups, no two cases represented the turn-behavior precisely at the same time or with the same intensity. We discuss three different potential explanations:

Spatial Environments: Each route was walked twice by one familiar and one unfamiliar person, i.e., there were 16 different routes considered in this study. Differences in within-class cases might, hence, stem from the environmental differences among the routes, e.g., urban configuration, segment (e.g., length) and junction characteristics e.g., number of segments, etc.), POIs, etc.

Levels of Familiarity: Theoretical reasoning [28, 10] suggests that there are different levels of spatial familiarity. Consequently, differences in turning-behavior during the matching-to-action phase may be considered as a potential indicator of different levels of familiarity. However, collecting ground truth data for these levels is still an open research question. Researchers tend to use either the number of years a person has lived in a city as an indicator (see e.g. [30]) or collect level of familiarity data using custom-designed questionnaires (see, e.g. [25] who collected self-report familiarity ratings on a 7-point scale but re-classed it to binary measure for their analysis.)

Users' Spatial Abilities: A considerable body of literature stresses the importance of individuals' spatial abilities in all spatial tasks including wayfinding [29, 16]. Thus, another possible explanation for the difference in turning activity could be individual differences in spatial abilities, which could act as a moderator for the effects of spatial familiarity.

All of these explanations can be valid and the differences observed within each class may stem from one or many of these factors. However, we have slight evidence that highlights the priority of *Levels of Familiarity* more than others. As presented in Figure 4, even for the single case in the unfamiliar class for which two different wayfinders walked the same route, we can observe these individual variations in turning-behaviors. Although this one case is not considered as a representative sample, observing a case like this, with a fixed spatial environment for both wayfinders, reinforces the finding that the variations in turn-behaviors within each class may stem from the levels of familiarity (see Future Work below).

6 Conclusion and Future Work

In this paper, we provide evidence that familiarity of wayfinders can be detected by analyzing their gaze behavior during the matching-to-action phase of decision-making for turning activity. We draw this conclusion based on the analysis of gaze data that has been collected during an in-situ wayfinding experiment by a customized pre-trained XGBoost turning activity classifier. The classification results indicate a distinguishable pattern between these two groups regarding their turning-behavior. Within each group, however, we also observe unique patterns of turning-behavior. We discussed this observation with respect to three possible explanations: the impact of the spatial environment, different levels of familiarity, and users' spatial abilities. Each of these factors may account for the within-group differences in observed turn-behavior. However, a single-case observation hints that spatial environment may not be the most important factor, as in this case, two different wayfinders, both unfamiliar, walked the same route. Hence, the results of the current study open the door to predicting, modeling, and hopefully defining spatial familiarity on a continuous scale. This leaves room for further investigations (in both controlled and uncontrolled settings) concerning all of these factors in general and different levels of familiarity in particular. This investigation will be fostered by the fact that spatial ability and spatial environment can both be fairly controlled in experimental designs. For instance, to disentangle the user and environmental effects, it would be interesting to conduct an experiment with relatively similar routes for each wayfinder, so that each wayfinder can be assigned a comparable familiar and unfamiliar route concerning the environmental factors. Such a setting allows for further within-class analysis, e.g., the gaze-pattern differences between familiar and unfamiliar routes walked by the same person. Another experiment for keeping the environment factor fixed would be to select a single route and recruit only familiar or unfamiliar participants. Adding more behavioral data sources may as well be worthwhile to consider in this research endeavor.

References

1. Linda P. Acredolo. The familiarity factor in spatial research. *New Directions for Child and Adolescent Development*, 1982(15):19–30, 1982.
2. N. Alinaghi, M. Kattenbeck, A. Golab, and I. Giannopoulos. Will you take this turn? gaze-based turning activity recognition during navigation. In *Proc. of the 11th Intl. Conf. on Geographic Information Science (GIScience 2021)-Part II*, 2021.
3. G. L Allen. Spatial abilities, cognitive maps, and wayfinding. *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, 4680, 1999.

4 A. Bulling, J. A Ward, H. Gellersen, and G. Tröster. Eye movement analysis for activity recognition using electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):741–753, 2010.

5 G. T. Buswell. *How people look at pictures: a study of the psychology and perception in art.* Univ. Chicago Press, 1935.

6 H. Couclelis, R. G Golledge, N. Gale, and W. Tobler. Exploring the anchor-point hypothesis of spatial cognition. *J.of Environmental Psychology*, 7(2):99–122, 1987.

7 A. Dehghani, O. Sarbishei, T. Glatard, and E. Shihab. A quantitative comparison of overlapping and non-overlapping sliding windows for human activity recognition using inertial sensors. *Sensors*, 19(22):5026, 2019.

8 W. Dong, H. Liao, B. Liu, Z. Zhan, H. Liu, L. Meng, and Y. Liu. Comparing pedestrians' gaze behavior in desktop and in real Envs. *Cartography & GIScience*, 47(5):432–451, 2020.

9 C. Galdi, H. Wechsler, V. Cantoni, M. Porta, and M. Nappi. Towards demographic categorization using gaze analysis. *Pattern Recognition Letters*, 82:226–231, 2016.

10 N. Gale, R. G Golledge, W. C Halperin, and H. Couclelis. Exploring spatial familiarity. *The Professional Geographer*, 42(3):299–313, 1990.

11 I. Giannopoulos. *Supporting Wayfinding Through Mobile Gaze-Based Interaction.* PhD thesis, ETH Zurich, 2016.

12 L. Gokl, M. Mc Cutchan, B. Mazurkiewicz, P. Fogliaroni, and I Giannopoulos. Towards urban environment familiarity prediction. *Adv. in Cartography and GIScience of the ICA*, 2(5), 2019.

13 Antonia Golab, Markus Kattenbeck, Georgios Sarlas, and Ioannis Giannopoulos. It's also about timing! when do pedestrians want to receive navigation instructions. *Spatial Cognition & Computation*, pages 1–33, 2021.

14 R. G Golledge. Place recognition and wayfinding: Making sense of space. *Geoforum*, 23(2):199–214, 1992.

15 I. M Harms, B. RD Burdett, and S. G Charlton. The role of route familiarity in traffic participants' behaviour and transport psychology research: A systematic review. *Transportation Research Interdisciplinary Perspectives*, 9:100331, 2021.

16 M. Hegarty and D. Waller. *Individual differences in spatial abilities.* Cam. Univ. Press, 2005.

17 J. M Henderson, S. V Shinkareva, J. Wang, S. G Luke, and J. Olejarczyk. Predicting cognitive state from eye movements. *PloS One*, 8(5):e64937, 2013.

18 C. Hölscher, T. Meilinger, G. Vrachliotis, M. Brösamle, and M. Knauff. Up the down staircase: Wayfinding strategies in multi-level buildings. *J. of Env. Psychology*, 26(4):284–299, 2006.

19 T. Ishikawa and D. R Montello. Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, 52(2):93–129, 2006.

20 P. Kiefer, I. Giannopoulos, and M. Raubal. Using eye movements to recognize activities on cartographic maps. *Proc of SIGSPATIAL 2013*, pages 478–481, 2013.

21 P. Kiefer, I. Giannopoulos, M. Raubal, and A. Duchowski. Eye tracking for spatial research: Cognition, computation, challenges. *Spatial Cognition & Computation*, 17(1-2):1–19, 2017.

22 R. V Levine and A. Norenzayan. The pace of life in 31 countries. *J.of Cross-Cultural Psychology*, 30(2):178–205, 1999.

23 R. Li and A. Klippel. Wayfinding behaviors in complex buildings: The impact of environmental legibility and familiarity. *Environment and Behavior*, 48(3):482–510, 2016.

24 H. Liao, W. Dong, H. Huang, G. Gartner, and H. Liu. Inferring user tasks in pedestrian navigation from eye movement data in real-world environments. *International J.of Geographical Information Science*, 33(4):739–763, 2019.

25 H. Liao, W. Zhao, C. Zhang, W. Dong, and H. Huang. Detecting individuals' spatial familiarity with urban environments using eye movement data. *CEUS*, 93:101758, 2022.

26 B. Liu, W. Dong, Z. Zhan, S. Wang, and L. Meng. Differences in the gaze behaviours of pedestrians navigating between regular and irregular road patterns. *ISPRS Int J.of Geo-Information*, 9(1), 2020.

27 L. Mathur, M. Spitale, H. Xi, J. Li, and M. J Matarić. Modeling user empathy elicited by a robot storyteller. In *9th Intl. Conf. on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8, 2021.

28 F. Meneses and A. Moreira. Using gsm cellid positioning for place discovering. In *2006 Pervasive Health Conf. and Workshops*, pages 1–8. IEEE, 2006.

29 D. R Montello, K. L Lovelace, R. G Golledge, and C. M Self. Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American geographers*, 89(3):515–534, 1999.

30 Raffaella Nori and Laura Piccardi. I believe I'm good at orienting myself... But is that true? *Cognitive Processing*, 16(3):301–307, 2015.

31 M. J O'Neill. Effects of familiarity and plan complexity on wayfinding in simulated buildings. *J.of Environmental Psychology*, 12(4):319–327, 1992.

32 J. Platt et al. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers*, 10(3):61–74, 1999.

33 A. Rousell and A. Zipf. Towards a landmark-based pedestrian navigation service using OSM data. *ISPRS Int J.of Geo-Information*, 6(3), 2017.

34 N Savage et al. Seems familiar: An algorithm for inferring spatial familiarity automatically, accessed feb. 5, 2013. *Computer Science Department, University of California, Santa Barbara, CA*, 2013.

35 V. Schnitzler, I. Giannopoulos, C. Hölscher, and I. Barisic. The interplay of pedestrian navigation, wayfinding devices, and Enval features in indoor settings. In *Proc of ETRA 2016*, pages 85–93, 2016.

36 Y. Shiga, T. Toyama, Y. Utsumi, K. Kise, and A. Dengel. Daily activity recognition combining gaze motion and visual features. *Adjunct Proc of UbiComp 2014*, pages 1103–1111, 2014.

37 M. van Haeren and W. Mackaness. The influence of familiarity on route choice: Edinburgh as a case study. In *Proc. of GISRUK Conf., Leeds*, 2015.

38 W. Wang. Using location-based social media for ranking individual familiarity with places: a case study with foursquare check-in data. In *Progress in Location-Based Services 2014*, pages 171–183. Springer, 2015.

39 F. Wenczel, L. Hepperle, and R. von Stülpnagel. Gaze behavior during incidental and intentional navigation in an outdoor Env. *Spatial Cognition & Comp.*, 17(1-2):121–142, 2017.

40 J. M Wiener, S. J Büchner, and C. Hölscher. Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation*, 9(2):152–165, 2009.

41 A. L. Yarbus. *Eye movements and vision.* Springer, 2013.

42 W. Zhang, X. Zhao, and Z. Li. A comprehensive study of smartphone-based indoor activity recognition via xgboost. *IEEE Access*, 7:80027–80042, 2019.

43 Z. Zhou, R. Weibel, and H. Huang. Familiarity-dependent computational modelling of indoor landmark selection for route communication: a ranking approach. *International J.of Geographical Information Science*, pages 1–33, 2021.