

Article

Research on Generating an Indoor Landmark Saliency Model for Self-Location and Spatial Orientation from Eye-Tracking Data

Chengshun Wang ¹, Yufen Chen ^{1,*}, Shulei Zheng ¹, Yecheng Yuan ² and Shuang Wang ³

¹ Information Engineering University, Zhengzhou 450052, China; wcs2000-1@163.com (C.W.); z_score2@163.com (S.Z.)

² State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; yuanyc@reis.ac.cn

³ Sub-Institute of High-tech Standardization, China National Institute of Standardization, Beijing 100191, China; wangshuang@cnis.ac.cn

* Correspondence: cyfbeijing@163.com; Tel.: +86-371-8162-2294

Received: 26 November 2019; Accepted: 27 January 2020; Published: 4 February 2020



Abstract: Landmarks play an essential role in wayfinding and are closely related to cognitive processes. Eye-tracking data contain massive amounts of information that can be applied to discover the cognitive behaviors during wayfinding; however, little attention has been paid to applying such data to calculating landmark saliency models. This study proposes a method for constructing an indoor landmark saliency model based on eye-tracking data. First, eye-tracking data are taken to calculate landmark saliency for self-location and spatial orientation tasks through partial least squares regression (PLSR). Then, indoor landmark saliency attractiveness (visual, semantic and structural) is selected and trained by landmark saliency based on the eye-tracking data. Lastly, the indoor landmark saliency model is generated by landmark saliency attractiveness. Recruiting 32 participants, we designed a laboratory eye-tracking experiment to construct and test the model. Finding 1 proves that our eye-tracking data-based modelling method is more accurate than current weighting methods. Finding 2 shows that significant differences in landmark saliency occur between two tasks; thus, it is necessary to generate a landmark saliency model for different tasks. Our results can contribute to providing indoor maps for different tasks.

Keywords: indoor landmark; eye-tracking data; landmark saliency model; self-location; orientation

1. Introduction

Wayfinding to a destination through an indoor or outdoor environment is a purposive, directed, and motivated behavior for efficiently finding one's way [1,2]. Wayfinding also involves a series of challenging behaviors that require participants to be aware of their self-location and to orient themselves [3] with the assistance of representative sensory cues from the external environment. Landmarks play an important role in providing guiding information for wayfinding in the physical environment [4,5] and can accelerate decision-making processes, especially at decision points for changing direction. Albrecht [6] discovered that landmarks had a strong relationship with participants' spatial cognition and memory. Clearly, landmarks play an essential role as wayfinding enhancers and navigational error reducers and can affect wayfinding tactics and strategies.

Given the importance of landmarks, it is necessary to measure the saliency of different kinds of landmark. Raubal and Winter [7] proposed the first approach to automatically identifying landmarks and calculating landmark saliency. They defined three different kinds of landmark saliency: visual, semantic, and structural. For example, geographic objects are visually attractive if they are in

sharp contrast to their surroundings. Subsequently, researchers have gradually proposed various methods for analyzing landmark attractiveness, such as expert knowledge [8], eye-tracking [9] and electroencephalography (EGG) [10]. Among these methods, eye movements directly reflect users' visual behaviors and precisely measure landmark attractiveness. Thus, studies that involve inferring landmark salience from eye movements have emerged. To date, studies have explored the use of eye-tracking methods to compare the cognitive differences in navigation [11–13]. Although Jia [14] proposed a landmark salience model calculated by eye-tracking data, they only considered visual attractiveness, without semantic and structural attractiveness. In addition, various tasks can lead to the production of eye-tracking data that are different from those observed in relation to landmark salience. Researchers [11,15] have inferred wayfinding tasks by using eye-tracking data, but they did not figure out the relationship between tasks and landmark salience. Therefore, exploratory research on how eye-tracking data calculate the visual, semantic and structural salience of landmarks for different tasks is still a challenging endeavor.

In this article, qualifying eye-tracking data and accurate algorithms were selected to calculate indoor landmark salience. Due to self-location and orientation being critical tasks during wayfinding [11], the task differences in indoor landmark salience were compared between these two tasks. If the significant task difference occurs, then an indoor landmark salience model for self-location and orientation can be established. Specifically, we focused on two questions:

1. Can eye-tracking data be used to construct an indoor landmark salience model? If so, how can the accuracy of the salience results be ensured?
2. Are there any differences in landmark salience between self-location and orientation in indoor wayfinding? If differences occur, how can an indoor landmark salience model be built for self-location and orientation?

This study makes two main contributions. On the one hand, our feature selection method and weighting algorithm are beneficial for understanding the relationship between eye movement metrics and indoor landmark salience and can extend the calculation method for indoor landmark salience. On the other hand, comparing the differences in landmark salience between self-location and orientation is also helpful for researchers to redesign different indoor landmarks on navigation maps for various wayfinding tasks.

The rest of this article is organized as follows. The related work is presented in Section 2. Section 3 presents the method used to construct the indoor landmark salience model. Section 4, a case study, is designed to test the model and compare it with the differences in landmark salience in two tasks. Section 5 discusses the important factors for the construction of the landmark salience model and compares it with previous studies. Section 6 ends this report with a conclusion and directions for future research.

2. Background and Related Work

2.1. Indoor Landmark Salience Models

Landmarks are important features in route directions during wayfinding. Sorrows and Hirtle [16] defined landmarks as prominent objects that individuals use as a reference point to help them in memorizing and recognizing routes, as well as locating themselves in terms of their ultimate destination. The aim of landmark identification is to find all the geographic objects in a given region that may in principle serve as a landmark [17]. To quantitatively compute landmarks, the concept of landmark salience has been proposed. Landmark salience is based on the concept of attractiveness, which reflects the importance of each landmark. The generation of a landmark salience model includes two major components, landmark salience attractiveness and weighting methods.

On the one hand, Raubal and Winter [7] presented the first approach to classifying landmark salience attractiveness, dividing landmark salience into three types of attractiveness (visual, semantic

and structural salience) to identify landmarks. Based on this finding, Elias [18] introduced building labels, building density and road orientation to describe the salience of geographic objects. Richter and Winter [17] defined the formal model for landmark salience, which includes four measures of visual attractiveness: the façade area, shape, color, and visibility. Zhu [19] proposed the façade area, the board size and design features to calculate the salience of indoor landmarks. However, there is no consensus regarding the salience attractiveness classification of geographic objects.

On the other hand, it is necessary to weight salience attractiveness. Currently, weighting methods, such as questionnaires, documentary sources and expert knowledge, are used to measure landmark salience. Mummidi and Krumm [20] calculated salience by comparing the number of times a specific n-gram appears in a cluster (term frequency) with the number of times the same n-gram appears in all clusters combined (document frequency). Wang [21] combined expert knowledge and proposed some definitions from cognitive and computational perspectives to evaluate indoor landmark salience. However, such methods are cumbersome and labor intensive. Furthermore, the results of such methods depend on the available data but will fail to detect other data because there are only very few data.

In addition, recent years have seen rapid advances in indoor spatial data modelling and an increasing availability of indoor geographic information system (GIS) data [22]. As a result of these rapid advances in indoor data modelling, many innovative indoor location-based service (LBS) applications have been developed, such as indoor wayfinding [23]. Thus, the indoor landmark salience models have been researched in recent years. Researchers [14,19] have proposed indoor landmark salience models based on visual, semantic and structural attractiveness, which is similar to the formal outdoor salience model. However, these attractiveness parameters (visual, semantic and structural) in the outdoor salience model cannot be directly applied to the indoor salience model. On the one hand, the landmark attractiveness factors in indoor spaces differ from those in outdoor spaces. Although Li L [24] has proposed that outdoor landmark attractiveness (shape factor, color and size) can be applied to describe indoor landmark salience, the cultural and historical importance in outdoor landmark attractiveness cannot be directly taken to describe landmarks in indoor environments, such as malls or airports. On the other hand, the spatial arrangement of indoor spaces differs from that of outdoor spaces. For example, multiple kinds of object can be regarded as landmarks in outdoor environments [25], such as churches, shopping malls, and bridges. However, these objects cannot be regarded as landmarks in indoor environments [26]. Lyu [27] mentioned that landmarks can be classified into four types: architecture (pillars and fronts), function (doors, stairs, and elevators), information (signs and posters) and furniture (tables, chairs, benches and vending machines). Thus, it is essential to propose a landmark salience model for indoor environments.

2.2. Differences in Landmark Salience during Wayfinding

The current progress in the cognitive sciences relevant to wayfinding investigates how to identify relevant landmarks, how to improve route instructions, and how to compute a better route [3]. Landmarks at decision points are important features in route directions during wayfinding. However, there are a large number of possible landmarks that can be included in route instructions in different situations and for different travelers [28]. Different travelers will find different landmarks to be most useful in a given situation.

There are three important dimensions (personal, navigation system-related and environmental) that impact the differences in landmark salience in wayfinding [29]. Among these dimensions, the personal dimension plays the most important role in person-centric navigation, and it is in this dimension that the most differences in wayfinding occur. For instance, Nuhn [8] identified the personal dimension and its attributes by taking into account five dimensions: personal knowledge, personal interests, personal goals, personal background and individual traits. This author proposed a personal landmark salience model based on these dimensions. In addition, the dimensions of a wayfinding task emerge based on landmark differences. Although task inference has been widely researched in

pedestrian navigation, few researchers have further investigated the task dimensions in landmark salience, especially the landmark differences in person-centric wayfinding.

Wayfinding is a cognitive behavior for finding a distal destination with a series of tactical and strategic tasks [30], including reading a map, remembering the route, finding one's location and maintaining one's orientation with external features or landmarks. Various tasks result in different forms of visual attention and cognitive behaviors with regard to landmarks. Thus, landmark salience keeps changing as participants accomplish different tasks. Two crucial behaviors during wayfinding are self-location and spatial orientation [31]. In self-location, one identifies his or her position in a spatial setting, and it includes several sub-processes, such as map orientation, feature matching, and configuration matching [17]. Spatial orientation is closely related to self-location; it involves determining the direction that one is facing when given an external instruction (cognitive or real maps) [32]. Wiener [33] reported a gaze bias between free exploration and pre-set route tasks. Participants displayed a significant tendency to choose the path leg offering the longest line of sight during free exploration, but that trend did not occur in the chosen route group. Wang [21] reported that although males and females had similar levels of effectiveness and efficiency in self-location, route memorization, and route following, there was a significant difference between them in map reading and indoor wayfinding tasks. However, little research has measured the differences in landmark salience between the self-location and orientation tasks in indoor environments.

2.3. Eye-Tracking for Task Differences in Landmark Salience

There are two important factors in comparing the task differences in landmark salience models: landmark salience calculation and statistical analyses of task differences. Researchers have adopted multiple methods [34–36] to calculate landmarks during wayfinding, such as questionnaires, pose estimation, and eye-tracking methods. Among these methods, eye-tracking can directly capture visual attention to landmarks in a quantitative way. Both quantitative and qualitative analyses of eye-tracking data can be applied to determine whether significant differences in salience occur during wayfinding.

On the one hand, in recent years, the eye-tracking method has been gradually proposed to analyze spatial cognitive performance in landmark identification because the user's gaze can provide an easy, fast and natural way to capture visual behaviors with regard to landmarks [37]. In addition, eye-tracking data assist researchers in analyzing gaze performance in quantitative ways [38]. For instance, eye-tracking data have been used as a rich data source for mining visual attention to landmarks during wayfinding [39,40]. Only recently have eye-tracking data been used by Jia [14] to calculate the visual attractiveness of landmarks. However, this author only resolves the problem of calculating the visual salience model. The use of eye-tracking data to generate a landmark salience model (visual, semantic and structural attractiveness) has not been tested. Researchers have applied the eye-tracking method to analyze semantic and spatial information [21,40]. For example, Raubal [7] proposed that city maps and street graphs are complemented with images and content databases, which could provide visual data as well as semantic and structural data. Wang [21] extracted areas of interest (AOIs) on indoor maps to analyze the semantic information of landmarks. Kiefer [41] proposed that the eye-tracking method could be taken to analyze route information and landmarks at decision positions, reflecting the structural attractiveness of landmarks. Thus, it is theoretically possible to calculate landmark semantic and structural salience model by eye-tracking data.

On the other hand, participants produce different eye movement patterns with regard to landmarks as their tasks change [41,42]. Kiefer [43] applied machine learning methods to detect six common map activities from eye-tracking data, proving that eye-tracking data can be applied to distinguish user tasks. Liao [15] demonstrated that wayfinding tasks (self-location, orientation, route remembering) can be inferred by eye-tracking data in outdoor environments, opening the door to potential indoor wayfinding applications that can provide task-related information depending on the task that a user is performing. However, little research has measured the differences in landmark salience between self-location and orientation based on eye-tracking data.

3. Indoor Landmark Salience Model

Based on the discussion of the related work, eye-tracking data can be applied to measure landmark salience. The core of this calculation method consists of regarding eye-tracking data as a mediator that is taken to measure the coefficient of landmark salience attractiveness (visual, semantic, structural). For example, Jia [14] proposed a visual salience model of landmarks based on eye-tracking experiments. The key to this author's method consisted of using eye-tracking data to represent a salience model with controllable and computable visual attractiveness, which is essential to improve the accuracy of the salience model and to provide a solution that addresses the landmark discrimination in interactions between humans and environments. However, the author did not consider structural or semantical attractiveness, and the landmark differences between various tasks were not discussed. In this section, we propose a method to generate an indoor landmark salience model based on eye-tracking data that considers the differences in self-location and orientation tasks.

3.1. Landmark Salience Based on Eye-Tracking Data

We defined landmark salience based on eye-tracking data as the salience results of objects calculated by eye-tracking data. Landmark salience based on eye-tracking data S_{eye} was calculated based on the stimulated landmark salience (S_{sti}), while the accuracy of S_{eye} was determined in two ways, by selecting the eye-tracking data and by calculating the coefficients of eye-tracking data.

3.1.1. Stimulated Landmark Salience

The concept of stimulated landmark salience S_{sti} was introduced by Jia [14]. S_{sti} means stimulated landmark salience results, which is measured by the percentage of participants who selected the object as a landmark in one specific setting [14]. In this paper, we selected eight indoor scene images as the specific settings. Participants were required to observe these images and select their favourite landmark in each of these images. The most attractive landmarks in each of these images were selected to calculate the stimulated landmark salience (S_{sti}). The percentage of participants who chose the most attractive landmark was measured as the result of S_{sti} . S_{sti} played an important role in calculating S_{eye} . On the one hand, eye-tracking data that have no statistically significant relationship with S_{sti} was deleted; on the other hand, S_{sti} was taken to measure the coefficient of eye-tracking data.

3.1.2. Eye-Tracking Data Selection

1. Data classification

Based on previous studies, eye-tracking data (fixations, saccades and pupil) have been widely adopted in eye-tracking studies [15,44]. The quantitative analysis of visual search strategies was closely related to eye-tracking data in predetermined areas of interest (AOIs) [45]. Thus, eye-tracking data are collected in both AOIs and the total area. The description of each example of eye-tracking data is provided in Table 1.

Table 1. Type of eye-tracking data.

Type	Features	Unit Statistic	Variable Definitions
Fixation			
Total	Total fixation duration	Second (s)	The total duration of fixations
	Total fixation counts	Count	The total counts of fixations
	Total fixation dispersion	pixel	The total dispersion of fixations
AOI	Time to first fixation	Second (s)	The time before first fixation on AOIs
	First fixation duration	Second (s)	The duration first fixation taken on AOIs
	Gaze duration	Second (s)	The duration of fixations on AOIs
	Fixation dispersion	pixel	The dispersion of fixations on AOIs
	Fixation counts	Count	The counts of fixations on AOIs

Table 1. Cont.

Type	Features	Unit Statistic	Variable Definitions
Saccade			
Total	Total saccade counts	Count	The total counts of saccades
	Total saccade duration	Second (s)	The total duration of saccades
	Saccade amplitude	Degree (°)	The total amplitude of saccades
AOI	Saccade counts	Count	The counts of saccades on AOIs
	Saccade duration	Second (s)	The duration of saccades on AOIs
	Saccade amplitude	Degree (°)	The amplitude of saccades on AOIs
Pupil			
Total	Pupil diameter	Millimeter (mm)	The average left and right pupil diameter
AOI	AOI pupil diameter	Millimeter (mm)	The average left and right pupil diameter on AOIs
	Pupil difference	Millimeter (mm)	The differences between Pupil and AOI pupil diameter

2. Data selection

Eye-tracking data that were significantly different from the stimulated visual salience S_{sti} were selected to calculate landmark salience.

• Normalization

Normalization can transform a dimensional expression into a dimensionless expression so that the indexes of different units or scales can be compared and weighted. The features are converted to a decimal value ranging from 0 to 1 through min-max normalization [46].

$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

$$x_{max} = \max_{1 \leq i \leq N} x_i, x_{min} = \min_{1 \leq i \leq N} x_i.$$

• Selection process

To avoid the uncertainty errors associated with landmark salience based on eye-tracking data, the normalized features should have a statistically significant relationship with the salience results. Stimulated landmark salience (S_{sti}) was regarded as the landmark salience result. Then, one-way ANOVA was used to measure the significant differences between eye-tracking data and stimulated landmark salience (S_{sti}). Only significant features ($p < 0.05$) were selected in this research.

3.1.3. Weighting Algorithms

1. Algorithm selection

The weighting algorithm is an essential method to calculate the correlation. To guarantee the reliability of correlation results, five weighting methods were used in our research. These methods include partial least square regression (PLQR), Analytic Hierarchy Process (AHP), Entropy weight method (EWM), Standard deviation method (SDM) and the Critic method. These five weighting methods are classic and commonly used. The weighting results were calculated by SPSS 11.0. The most accurate algorithm was selected in this paper.

2. Accuracy test

The precision of the weighting method was tested by the absolute difference value between stimulated salience (S_{sti}) and the visual salience calculated by eye-tracking data (S_{eye}). A smaller difference in the test results proved the better accuracy of the weighting method.

$$Accuracy = |S_{eye} - S_{sti}| \quad (2)$$

3.1.4. Calculating Process

Based on previous finding [14], S_{eye} was measured by the product of the eye-tracking data and their coefficients, and the S_{sti} was taken to calculate the coefficients of the eye-tracking data. Thus, the formula of landmark salience based on eye-tracking data S_{eye} is proposed as follows:

$$S_{eye}(x) = \sum_{i=1}^n \lambda_i e_i \sim S_{sti}(x) \quad (3)$$

where x is the name of a landmark, n represents the amount of eye-tracking data types, λ denotes different types of eye-tracking data, and e is the coefficients of eye-tracking data. The larger the value of e is, the greater the importance of λ .

The calculation process of landmark salience based on eye-tracking data is presented in Table 2.

Table 2. Calculation method for landmark salience based on eye-tracking data.

Input: eye-tracking data λ and stimulated landmark salience S_{sti}
Output: landmark salience based on eye-tracking data S_{eye}
for the result of each λ_i and S_{sti} calculated by one-way ANOVA $\in \beta$ do If $\beta_i < 0.05$ then return λ_i end If $\beta_i > 0.05$ then delete λ_i end end /* feature selection */ for the S_{sti} and S_{eye} calculated by the five weighting algorithms do calculate the absolute difference value between stimulated salience (S_{sti}) and the landmark salience based on eye-tracking data (S_{eye}); select the weighting algorithm that results in the lowest difference value as the most accurate algorithm; end /* weighting algorithm comparison */ for the selected weighting algorithms do calculate the coefficient of eye-tracking data (e), and establish the landmark salience based on eye-tracking data (S_{eye}). end /* landmark salience based on eye-tracking data */

3.2. Indoor Landmark Salience Model

3.2.1. Visual Attractiveness

1. Shape features

An outstanding shape is an essential salient attribute. According to the definition by Richter and Winter [17], the shape factor and deviation were selected. Put simply, the shape factor is the proportion of height and width. Deviation is the ratio of the area of the minimum-bounding rectangle (mbr) of the object's façade to its façade area [26]. Unusual shapes and deviations, especially among more regular, box-like objects, are highly remarkable.

2. Colour features

A landmark is salient if its colour or lightness contrasts with the surrounding objects. We used the hue error (Δh) and lightness to measure landmark colour. The hue error can be used to compare the hue value difference between a landmark and an indoor environment. The RGB of the landmark and floor was converted to LAB by Photoshop, and the hue errors were calculated by ColorTell

(www.colortell.com). Lightness was measured as though the landmark contained a bright section (window, door, pictures). If there was a bright section in the landmark, then the value of lightness was 1.

3. Façade area

The façade area is used to calculate the size of a landmark [26]. If the façade area of an object is significantly larger or smaller than the façade areas of the surrounding objects, then this object becomes well noticeable. The façade area was measured as height multiplied by width.

4. Visibility features

Visual distance is used to measure visibility. Clearly, if the visual distance of a landmark is shorter than that of other objects, then the landmark will be more noticeable. Visual distance was measured as the shortest distance between the participant's location and the landmark.

The detailed information of visual attractiveness was shown in Table 3.

Table 3. Visual attractiveness.

Indicator	Property	Unit Statistic	Measurement
Shape	shape factor *	Ratio	Shape factor = height/width
	deviation *	Ratio	Deviation = area mbr /Facade Area
Color	hue error **	Decimal	$(n - 1) \times 30^\circ \leq \Delta h \leq n \times 30^\circ \rightarrow 0.2 \times n$
	lightness **	Boolean value	If lightness $c = 1$; else $c = 0$
Façade area	Façade area *	Square meter (m ²)	Facade Area = height \times width
Visibility	Visual distance **	Meter (m)	Visual Distance = min{Perceivable distance}

* in table are referenced from [17], ** are referenced from [14].

3.2.2. Semantic Attractiveness

1. Semantic importance

This property reflects whether an object has an important meaning. Semantic importance represents the proportion of AOI fixation duration and total fixation duration during map reading tasks. The AOI includes the name and point symbol of an object on the map.

Just and Carpenter [47] mentioned that a longer fixation duration either means difficulty in understanding information or represents that the participants show more interest. However, the former explanation is rejected in this paper because participants were educated and can interpret the semantic information on the map. In addition, the participants were driven by the tasks to find and remember important landmarks on maps without a time limit. Thus, the longer the amount of time of visual attention is given to an AOI, the greater the semantic importance of the meaning of the object. Based on previous research [14,48], the AOIs include the name and symbol of objects with a buffer due to the imprecision in eye movements. The AOIs are shown in Figure 1.

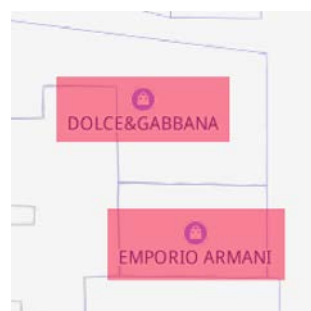


Figure 1. AOI divisions for semantic importance.

2. Explicit marks

An object may have explicit marks, such as signs on the front of a store. These signs explicitly label an object, communicating its semantics. This property was assessed by a Boolean value.

3. Degree of familiarity

This property indicates whether participants are familiar with a mark. First, the object must have a mark; otherwise, the value is 0. Then, the degree of familiarity was calculated by the proportion of participants who were familiar with the mark (Table 4).

Table 4. Semantic attractiveness.

Indicator	Unit Statistic	Measurement
Semantic importance	Ratio	AOI fixation duration/Total fixation duration
Explicit marks *	Boolean value	If explicit 1, or 0
Degree of familiarity	Ratio	Familiar with the mark/Total participants

* in table are referenced from [17], Unstarred characteristics and their measurements are mentioned by our research.

3.2.3. Structure Attractiveness

1. Number of adjacent routes

Objects located at intersections are more important for route instructions than objects located along routes. If an object is adjacent to more than one route, then it is located at a street intersection and is therefore more suitable. To assess landmark salience, the number of edges adjacent to the object was stored.

2. Number of adjacent objects

Freestanding objects are more attractive than objects with many neighbours. This attribute is mainly important for signs and elevators because other objects, such as stores, are normally connected to other structures. The number of adjacent objects was stored to assess structural salience.

3. Location importance

Location importance indicates the attractiveness of objects caused by different locations. Location importance can be calculated by the distances between one object and the nearest nodes. Nodes are the intersections in a network [17]. In this paper, intersections in the David Mall indoor map (map.baidu.com) were selected as nodes.

$$L(x) = 1/d(y, x) \quad (4)$$

where x is the node and y is an object. The $d(y, x)$ denotes the distance between the nodes x and the object y . Since no two rooms occupy the same space, the distance between any x and y will not be 0. If the distance between x and y is less than 1 m, then $L(x)$ is 1 (Table 5).

Table 5. Structural Attractiveness.

Indicator	Unit Statistic	Measurement
Number of adjacent routes *	Constant	The number of routes
Number of adjacent objects *	Constant	The number of objects
Location importance **	Ratio	$1/d(y, x)$

* in table are referenced from [49], ** in table are referenced from [27].

3.2.4. Modelling Process

1. Comparing the differences in landmark salience between self-location and orientation

It is essential to determine whether differences in landmark salience occur between self-location and orientation tasks before generating a landmark salience model for the two tasks. Landmark salience based on eye-tracking data S_{eye} was calculated for both self-location ($S_{eye}^{self-location}$) and orientation ($S_{eye}^{orientation}$). One-way ANOVA was used to measure the statistically significant differences between $S_{eye}^{self-location}$ and $S_{eye}^{orientation}$ using SPSS 11.0. The results ($p < 0.05$) indicate that a significant difference between self-location and orientation is found, and the landmark salience model can be generated for both tasks. Otherwise, it is meaningless to construct a visual salience model for the two tasks.

2. Indoor landmark salience model for two tasks

Jia [14] proposed that S_{eye} could be used to construct the coefficient of landmark salience attractiveness. With regard to whether a significant difference occurs between $S_{eye}^{self-location}$ and $S_{eye}^{orientation}$, the coefficients are different between the two tasks. Thus, $S_{landmark}^{tasks}$ represents the visual salience model for different tasks.

The formula of the visual salience model is as follows:

$$S_{visual}^{tasks}(x) = \sum_{i=1}^n f_i w_i \sim S_{eye}^{tasks}(x) \quad (5)$$

$$S_{semantic}^{tasks}(x) = \sum_{i=1}^n f_i w_i \sim S_{eye}^{tasks}(x) \quad (6)$$

$$S_{structural}^{tasks}(x) = \sum_{i=1}^n f_i w_i \sim S_{eye}^{tasks}(x) \quad (7)$$

$$S_{landmark}^{tasks}(x) = 1/3(S_{visual}^{tasks}(x) + S_{semantic}^{tasks}(x) + S_{structural}^{tasks}(x)) \quad (8)$$

where f denotes landmark attractiveness and w is the coefficient of landmark attractiveness. The tasks include self-location and orientation.

4. Case Study

4.1. Experimental Design

4.1.1. Apparatus

A Tobii X120 (Tobii AB, Sweden, www.tobii.com) eye tracker with a sampling rate of 120 Hz and a Samsung 22-inch monitor were selected. The X120 had a recording accuracy of 0.5° and might have a deviation of 0.1° . The spatial resolution was 0.3° , and the head movement error was within 0.2° . The visual tracking distance was between 50 and 80 cm. The monitor displayed the stimuli with a screen resolution of 1680×1050 pixels. Tobii Pro Analyzer software was used to manage and analyse the eye-tracking data. The research was capable of obtaining deep insights into visual saliency in regard to indoor pictures and maps by analysing eye-tracking data.

4.1.2. Procedure

The experiment was conducted in a quiet and well-lit room (Figure 2a). In the pre-test training, the participants were welcomed and were required to provide their personal information (gender, age, familiarity of the David Mall and experience using a computer in everyday life) and complete two skill tests. Previous research has reported that wayfinding tasks were examined in relation to spatial skills and self-reported skills [50]. The Mental Rotation Test (MRT) and the Santa Barbara Sense of Direction

Scale (SBSOD) could be used to test spatial skills and self-reported skills, respectively. The participants were asked to complete two tests before the experiment to ensure they had similar skills. In addition, as the Tobii X120 could not support the participants to check the stimuli again, the participants were informed to remember the experimental stimuli during the experimental procedure.



Figure 2. Eye-tracking experiments. (a) Experimental environment; (b) Number of experimental stimuli in the three tasks; (c) Visual area of the experimental stimuli in task 2; (d) Visual area of the experimental stimuli in task 3.

In the formal experiment, the participants were asked to complete three tasks (Figure 2b). The instructions for the participants are described below:

- Task #1 (landmark selection): Assume that you are shopping in a mall. When the experiment begins, you will view eight indoor scene images one at a time. Please select the most attractive landmark (store, bench, elevator or signs) in each image. When you find the result, click on the landmark to proceed. There is no time limit for you to find the landmark.
- Task #2 (self-localization): You will find your location on the map. First, you should observe an indoor scene image carefully and try to memorize the necessary landmark information as much as possible. You are not allowed to look at the image again. Then, you should find the location and click on it on the indoor map. Two locations should be found in this phase.
- Task #3 (orientation): You will find your orientation in the indoor scene image with the assistance of landmarks. You should remember the landmark information related to the route from A to B on the indoor map. Then, you will point out the correct orientation to get to B and click on it on the image. Two orientations need to be noted in this task. After that, the experiment ends.

4.1.3. Stimuli

Twelve panoramas were created as indoor scene images for the eye-tracking experiment. The panoramas were photographed using a Canon 800D camera with an 18–55 mm lens in Zhengzhou David Mall, China. The camera was fixed on a 1.5-m tripod. We took 60 pictures at each location. The computer-generated panoramas were made using PTGui (www.ptgui.com). However, 360° panoramas were not used because image distortions occurred as the 360° panoramas were dragged. In the meantime, it was difficult for the participants to recognize detailed information in the 360° panorama shown in one picture due to the limitation of screen size. Thus, we selected half of the panorama with a 180° visual angle (Figure 2c,d).

Based on the experimental procedure, eight indoor scene images were selected in task 1 (Figure 3a–h). When the participants observed one image, they clicked on their favorite landmark in the stimuli. After the experiment, the most attractive landmarks, highlighted in Figure 3a–h, were

defined as AOIs for analyzing the eye-tracking data. The AOIs were divided by research though Tobii Analyzer, which could be used to collect and calculate eye-tracking data within or without the AOIs. The display of these eight images followed the Latin square principle.

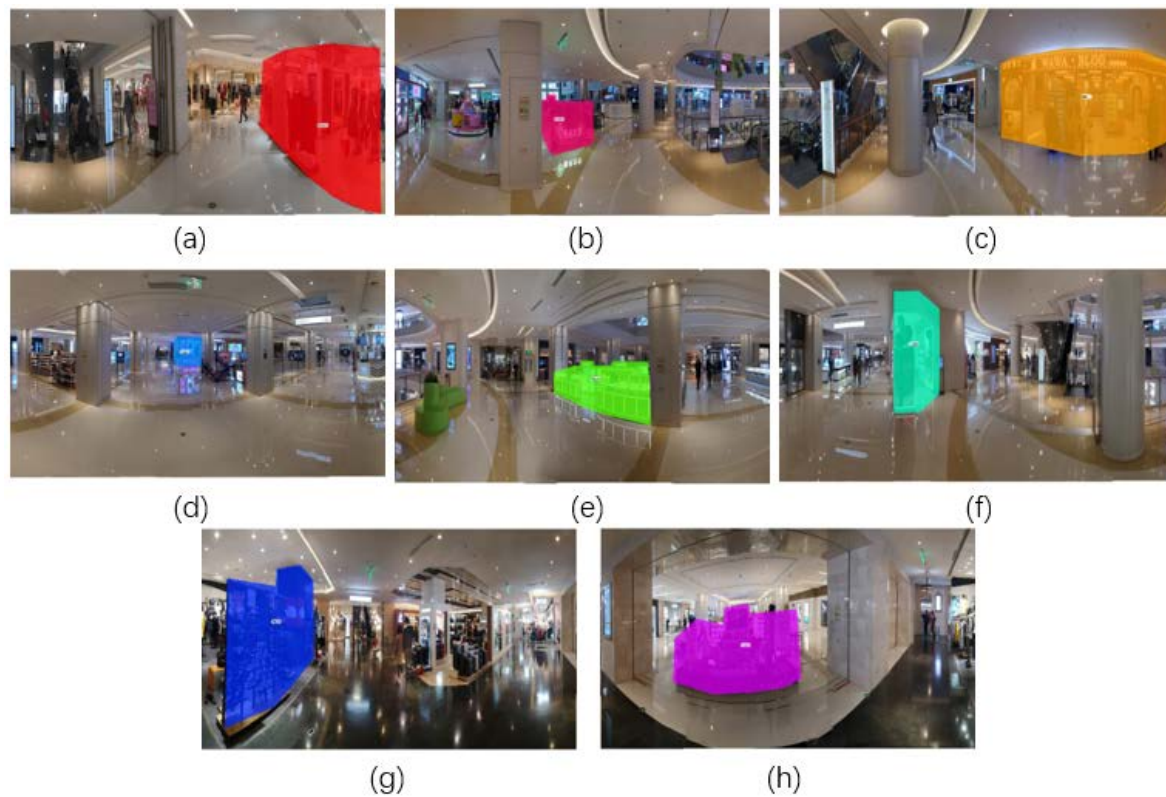


Figure 3. Eight images of the experimental stimuli in task 1. (a–h) are experimental stimuli in task 1.

In task 2, four Baidu indoor maps (map.baidu.com) were selected as indoor 2D maps. Baidu indoor maps are widely used by the general public in China, which ensures that the participants have similar levels of familiarity. Photoshop was applied to re-mark all of the landmarks and to redesign the point symbol in the same pattern (Figure 4(a2,b2,c2,d2)). To observe the orientation behaviors in task 3, navigation routes from A to B were drawn in these redesigned maps (Figure 4(a3,b3,c3,d3)).



Figure 4. Cont.



Figure 4. Experimental stimuli in task 3. (a1,b1,c1,d1) Indoor scene image; (a2,b2,c2,d2) Indoor map for self-location; (a3,b3,c3,d3) Indoor map for orientation.

The participants were required to view images to find their locations in task 2 and to remember the navigation routes to find their orientations in task 3. In order to avoid the participants observing the same indoor scene images in task 2 and 3, the participants were divided into two groups. The order of experimental stimuli was different between Groups 1 and 2. For the participants in Group 1, the order was Figure 4 a2-a1-b2-b1 in task 2 and Figure 4 c1-c3-d1-d3 in task 3. For Group 2, the order was Figure 4 c2-c1-d2-d1 in task 2 and Figure 4 a1-a3-b1-b3 in task 3.

4.1.4. Participants

A total of forty-six young male students majoring in cartography were recruited to join our pilot experiment as an experimental lesson. The results of five participants were omitted because their sample rates (calculated by Tobii Analyser) were below 80% [51]. Four participants were omitted because they did not pass SBSOD and MRT. Five participants did not continue to conduct the experiment because they were familiar with the David Mall. Thus, thirty-two participants continued to conduct the formal experiments. According to the experimental stimuli, the participants were divided into two groups. The sixteen participants in Group 1 were aged between 18 and 29 years old (mean age = 23.97, SD = 1.54). In Group 2, the sixteen participants were aged between 18 and 27 years old (mean age = 22.63, SD = 1.67).

All of the participants were familiar with computing technology. They all had normal or corrected-to-normal vision and could complete the experiment independently. The experiment was reviewed and approved by the local institutional review board (IRB). All of the participants provided their written informed consent to participate in the experiment.

4.2. Results

4.2.1. Landmark Saliency Based on Eye-Tracking Data

To answer question 1, with regard to whether eye-tracking data can be used to calculate landmark saliency, eye-tracking data were selected and weighted using five algorithms. According to the participants' selection, the stimulated landmark saliency (S_{sti}) results were 0.594, 0.906, 0.750, 0.594, 0.875, 0.813, 0.688 and 0.875 in Figure 3a–h, respectively.

1. Feature selection

All of the images (Figure 3a–h) in task 1 were used for feature selection. One-way ANOVA was used to compare the significance relationships between stimulated landmark salience (S_{sti}) and the eye-tracking data (S_{eye}); the results are shown in Table 6. Clearly, seven features have significant differences with stimulated landmark salience. These features, including total fixation duration, total fixation counts, gaze duration, fixation counts, total saccade duration, saccade duration and pupil difference, were selected to calculate the visual salience based on eye-tracking data. These features included fixation, saccade and pupil type, ensuring that all types of eye metric were considered in our research.

Table 6. Eye movement feature selection.

Type	Features	Mean	SD	ANOVA	
Fixation				F	p
Total	total fixation duration (s)	4.861	1.010	5.135	0.032 *
	total fixation counts	15.332	2.548	9.002	0.015 *
	total fixation dispersion	2105.042	716.124	2.411	0.123
AOI	time to first fixation (s)	0.612	0.353	1.277	0.277
	first fixation duration (s)	0.440	0.124	0.156	0.635
	gaze duration (s)	2.637	0.723	3.713	0.041 *
	fixation dispersion	320.402	71.188	0.263	0.419
	fixation counts	5.945	1.432	6.578	0.014 *
Saccade					
Total	total saccade duration (s)	0.561	0.093	18.351	0.001 *
	total saccade counts	13.134	3.086	1.770	0.205
	saccade amplitude	81.199	19.145	1.754	0.265
AOI	saccade duration (s)	0.292	0.063	6.613	0.004 *
	saccade counts	7.059	3.067	0.971	0.252
	saccade amplitude	49.419	16.540	1.910	0.244
	regression	1.480	0.565	1.138	0.094
Pupil					
Total	pupil size	3.771	0.040	3.597	0.071
AOI	AOI pupil size	3.875	0.041	3.783	0.068
	pupil difference	0.157	0.040	8.310	0.001 *

* means $p < 0.05$.

2. Feature weighting

To build the visual salience formula based on eye-tracking data, the eye-tracking data in Figure 3a–f in task 1 were used to measure the coefficient. Both statistical regression and weighting methods could be used to calculate the coefficient. To choose the best method, SPSS 11.0 was applied to compare these weighting algorithms. The results are shown in Table 7.

Table 7. Weighting algorithm for eye-tracking data.

Eye-Tracking Data		PLSR	AHP	EWM	SDM	CRITIC
Fixation						
	total fixation duration λ_1	0.005	0.172	0.108	0.187	0.142
	total fixation counts λ_2	0.007	0.180	0.112	0.171	0.154
	gaze duration (s) λ_3	−0.034	0.088	0.206	0.107	0.135
	fixation counts λ_4	0.003	0.116	0.139	0.128	0.122
Saccade						
	total saccade duration λ_5	−0.212	0.137	0.166	0.143	0.150
	saccade duration λ_6	1.631	0.174	0.116	0.145	0.147
Pupil						
	pupil difference λ_7	1.348	0.132	0.152	0.117	0.18
Intercept	C	0.128				

3. Results accuracy

The eye-tracking data in Figure 3g,h were collected to test the accuracy of the weighting algorithm. Figure 5 shows the difference value (dv) results of the participants. Clearly, the difference value of the SDM is higher than that of the other algorithms from five participants ($dv = 5.86$) to 32 participants ($dv = 3.99$), proving that the SDM is the worst method for calculating the coefficient. All of the lowest difference value results are observed for PLSR, which shows that PLSR is the best weighting method in this experiment. This finding confirms the previous evidence showing that PLSR is the most accurate method for visual salience based on eye-tracking data, as proposed by Jia [14].

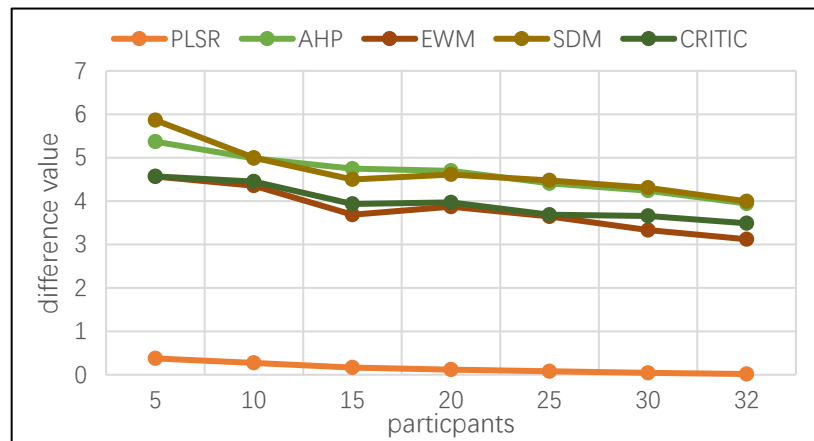


Figure 5. Accuracy of weighting algorithms.

4. Landmark salience based on eye-tracking data

Figure 5 shows that the different value results decrease as the number of participants among these algorithms increases. The average difference between 30 and 32 participants is 0.01, which shows that the number of participants is sufficient for this research. The formula of landmark salience based on eye-tracking data is shown as follows:

$$S_{eye}(x) = 0.005\lambda_1 + 0.007\lambda_2 - 0.034\lambda_3 + 0.003\lambda_4 - 0.212\lambda_5 + 1.631\lambda_6 + 1.349\lambda_7 + 0.128$$

4.2.2. Differences in Landmark Salience between Self-Location and Orientation

To answer question 2, with regard to whether differences in landmark salience occur between the self-location and orientation tasks, the visual salience of nineteen landmarks in tasks 2 and 3 was calculated based on eye-tracking data (formula in Section 3.1), and one-way ANOVA was applied to analyse the significant differences. The results are shown in Table 8.

Task 2 (self-location) was significantly different from task 3 (orientation) in landmark salience ($F = 4.156$ $p = 0.048 < 0.05$), indicating that the participants have a different visual performance regarding the AOIs in tasks 2 and 3.

Differences in landmark salience between tasks 2 and 3 are present in each AOI. According to Table 8, nineteen AOIs show significant differences in landmark salience between the self-location and orientation tasks. The participants in task 3 paid significantly greater visual attention to store AOIs (AOI1 = 1.105, AOI5 = 0.789, AOI6 = 1.287, AOI11 = 1.074, AOI14 = 1.148, AOI16 = 1.085 and AOI19 = 0.723) than did those in task 2 (AOI1 = 0.851, AOI5 = 0.413, AOI11 = 0.663, AOI14 = 0.835, AOI16 = 0.817 and AOI19 = 0.339). The elevator AOIs also show similar significant differences; the landmark salience of the elevator in task 3 is significantly higher than that in task 2, indicating that the participants paid more attention to store and elevator landmarks in the orientation task. However, it is difficult to determine the landmark salience of benches. Although the landmark salience of the bench in task 3 (AOI4 = 0.481) is significantly lower than those in task 2 (AOI4 = 0.346), those in AOI9 do not show the same tendency.

Table 8. Visual salience differences between task 2 and task 3.

AOI								
AOI Type	AOI1 store(cinema) *		AOI2 escalator *		AOI3 sign *		AOI4 bench *	
S_{eye}	T2: 0851	T3: 1.105	T2: 0.386	T3: 0.905	T2: 0.513	T3: 0.423	T2: 0.346	T3: 0.481
ANOVA	F = 32.152	$p = 0.001$	F = 9.401	$p = 0.004$	F = 24.152	$p = 0.001$	F = 78.921	$p = 0.001$
AOI								
AOI Type	AOI5 store(drinks) *		AOI6 store(eyeglass) *		AOI7 escalator*		AOI8 sign*	
S_{eye}	T2: 0.413	T3: 0.789	T2: 0.847	T3: 1.287	T2: 0.502	T3: 0.727	T2: 0.361	T3: 0.284
ANOVA	F = 42.115	$p = 0.001$	F = 15.083	$p = 0.001$	F = 12.502	$p = 0.001$	F = 8.928	$p = 0.006$
AOI								
AOI Type	AOI9 bench*		AOI10 store(drinks)		AOI11 store(Herborist) *		AOI12 store(cosmetic)	
S_{eye}	T2: 0.0.727	T3: 0.468	T2: 0.405	T3: 0.469	T2: 0.665	T3: 1.074	T2: 0.693	T3: 0.618
ANOVA	F = 16.024	$p = 0.001$	F = 3.152	$p = 0.032$	F = 109.293	$p = 0.001$	F = 1.526	$p = 0.423$
AOI								
AOI Type	AOI13 elevator		AOI14 store(jewelry) *		AOI15 escalator*		AOI16 store(Dior) *	
S_{eye}	T2: 0.501	T3: 0.618	T2: 0.835	T3: 1.148	T2: 0.212	T3: 0.355	T2: 0.817	T3: 1.085
ANOVA	F = 8.700	$p = 0.003$	F = 66.192	$p = 0.001$	F = 3.124	$p = 0.041$	F = 28.941	$p = 0.001$
AOI							ANOVA t2 and t3 dF (1,36) F: 4.156 p : 0.048	
AOI Type	AOI17 escalator		AOI18 store(D&G)		AOI19 store(coach) *			
S_{eye}	T2: 0.299	T3: 0.332	T2: 0.639	T3: 0.735	T2: 0.339	T3: 0.723		
ANOVA	F = 0.982	$p = 0.614$	F = 4.1865	$p = 0.052$	F = 34.512	$p = 0.001$		

* means $p < 0.05$.

4.2.3. Landmark Saliency Model for Self-Location and Orientation

The previous results in Section 4.2.2 show that differences in landmark saliency occur between tasks 2 and 3. Thus, the landmark saliency model can be constructed for the self-location and orientation tasks. Based on the modelling process, the landmark saliency attractiveness was normalized, and the results were compared with the landmark saliency based on eye-tracking data in tasks 2 and 3 using one-way ANOVA. Then, landmark saliency attractiveness with a significant difference was selected for regression through PLSR. The results are shown in Table 9.

Table 9 shows that the landmark saliency model includes nine factors for both tasks 2 and 3. The brightness, explicit marks and adjacent object factors were not selected for landmark saliency modelling. In addition, the coefficients of the selected factors were not the same. In task 2, the visual distance was significantly negatively correlated with S_{eye} . In task 3, the visual distance and degree of familiarity factors were negatively correlated with S_{eye} .

Table 9. The coefficients of landmark saliency attractiveness in tasks 2 and 3.

Measure	Property	Type	Task 2	ANOVA		Task 3	ANOVA	
			Coefficient	F	p	Coefficient	F	p
Visual	Shape factor	f_1	0.018	10.852	0.002 *	0.005	8.977	0.004 *
	Deviation	f_2	0.107	14.465	0.000 *	0.218	24.945	0.001 *
	Hue	f_3	0.060	12.388	0.001 *	0.159	23.355	0.001 *
	Brightness	f_4	—	0.081	778	—	1.146	0.291
	Façade area	f_5	0.0003	9.293	0.004 *	0.001	8.368	0.006 *
	Visual distance	f_6	−0.011	27.756	0.000 *	−0.014	26.769	0.000 *
Semantic	Intercept	b_1	0.559	—	—	0.706	—	—
	Semantic Importance	f_7	1.259	70.211	0.000 *	3.362	65.393	0.000 *
	Explicit marks	f_8	—	0.021	0.884	—	1.966	0.169
	Degree of familiarity	f_9	0.054	3.792	0.047 *	−0.098	10.829	0.002 *
Structural	intercept	b_2	0.368	—	—	0.332	—	—
	Adjacent routes	f_{10}	0.048	40.4	0.000 *	0.088	32.449	0.000 *
	Adjacent objects	f_{11}	—	3.4	0.073	—	1.244	0.272
	Location	f_{12}	0.228	8.124	0.007 *	0.339	5.298	0.021 *
	Intercept	b_3	0.283	—	—	0.288	—	—

* means $p < 0.05$.

According to Table 9, the landmark saliency models are generated as follows:

The landmark saliency model for self-location:

$$S_{visual}^{self-location}(x) = 0.018f_1 + 0.107f_2 + 0.060f_3 + 0.0003f_5 - 0.011f_6 + 1.259$$

$$S_{semantic}^{self-location}(x) = 1.259f_7 + 0.054f_9 + 0.368$$

$$S_{structural}^{self-location}(x) = 0.048f_{10} + 0.228f_{12} + 0.283$$

$$S_{landmark}^{self-location}(x) = 1/3(S_{visual}^{self-location}(x) + S_{semantic}^{self-location}(x) + S_{structural}^{self-location}(x))$$

The landmark saliency model for orientation:

$$S_{visual}^{orientation}(x) = 0.005f_1 + 0.218f_2 + 0.159f_3 + 0.001f_5 - 0.014f_6 + 0.706$$

$$S_{semantic}^{orientation}(x) = 3.362f_7 - 0.098f_9 + 0.332$$

$$S_{structural}^{orientation}(x) = 0.088f_{10} + 0.339f_{12} + 0.288$$

$$S_{landmark}^{orientation}(x) = 1/3(S_{visual}^{orientation}(x) + S_{semantic}^{orientation}(x) + S_{structural}^{orientation}(x))$$

5. Discussion

In this section, we first analyse the important factors that influence the generation of the indoor landmark salience model. We then discuss the differences in landmark salience between the self-location and orientation tasks from the perspective of the participants and indoor environments. Finally, we compare the accuracy of our model with previous weighting methods and propose improvements to our model.

5.1. Important Factors for the Landmark Salience Model

As indicated by the previous findings in Sections 4.2.1 and 4.2.2, question 1 with regard to whether landmark salience can be calculated by eye-tracking data has been answered. In this part, three important factors for the construction of the landmark salience model are shown.

The first factor is the type of eye-tracking data. To prove the reliability of the selected eye-tracking data (combined features), seven types of feature (combined features, fixation, saccade, pupil, fixation+saccade, fixation+pupil and saccade+pupil) were collected in the images of task 1, and PLSR was applied to calculate the coefficient. The difference value results between the stimulated visual salience and the predicted salience of the seven features are shown in Figure 6. Clearly, the combined feature has the lowest difference value (mean = 0.0022, SD = 0.0004), making it possible to improve the accuracy of visual salience based on fixation features and other types of eye-tracking data.

The second factor is the weighting algorithm. The previous results in Section 5.1 prove that PLSR is the best algorithm in this research. There are two reasons for this conclusion. On the one hand, the selected eye-tracking data are significantly different from the stimulated visual salience (Table 8), and the ANOVA results based on SPSS show that eye-tracking data follow a normal distribution, which means that the selected features could be used to establish multiple linear regression equations [14]. On the other hand, in PLSR, the stimulated visual salience and eye-tracking data are measured as dependent and independent variables, respectively, while the other algorithms consider only the variability of eye-tracking data.

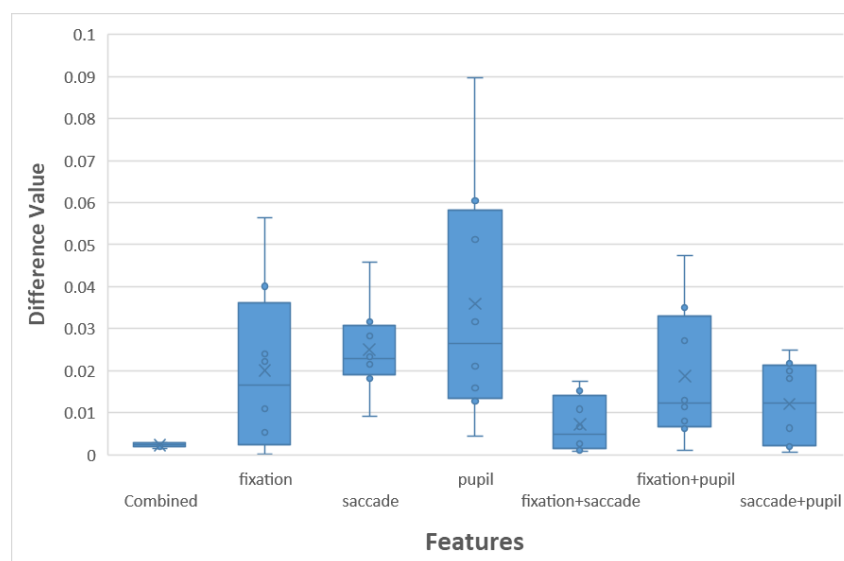


Figure 6. Comparison of mean difference value among eye-tracking data.

The last factor is the significance of landmark salience attractiveness. Table 9 reveals that the factor coefficients for the self-location and orientation tasks are different. For instance, the degree of

familiarity is significantly positively correlated in the self-location group, but it is negatively correlated with landmark salience based on eye-tracking data in the orientation group.

5.2. Differences in Landmark Salience between Self-Location and Orientation

Section 4.2 answered question 2 with regard to whether differences in landmark salience occurred between the self-location and orientation tasks. To explain this problem, the participants' visual behaviours and indoor environments are analysed in this section.

5.2.1. Differences in the Participants' Visual Behaviours

A t-test was applied to prove that the participants have significantly different visual behaviours in self-location and orientation (Table 10). There are significant differences between the two tasks in four types of eye-tracking data (total fixation duration, gaze duration, total saccade duration and saccade duration). For instance, the participants in the self-location group had a gaze duration of 1.124 s (SD = 0.535) on the self-location task, which is significantly less than that of the participants in the orientation task (mean = 1.818 s, SD = 0.789), indicating that the participants' gaze duration was shorter on the self-location task than on the orientation task.

Similar differences in total saccade features occur. The total saccade duration of the groups in tasks 2 and 3 was 1.059 s (SD = 0.176) and 1.597 s (SD = 0.126), respectively, which indicates that the participants in the self-location group had a significantly shorter saccade duration than did the orientation group ($t = -20.207$, $p < 0.001$). However, there are no significant differences in pupil features (pupil size, AOI pupil size and pupil difference) between the self-location and orientation groups, which indicates that the participants have similar pupil behaviours with regard to the AOIs in tasks 2 and 3.

Table 10. Differences in eye-tracking data between tasks 2 and 3.

Eye Movement Feature	Task2	Task3	t-Test	
	M ± SD	M ± SD	t	p
Fixation				
total fixation duration (s)	6.636 ± 0.955	10.694 ± 1.335	−16.089	0.001 *
first fixation duration (s)	0.335 ± 0.239	0.339 ± 0.230	−0.105	0.918
gaze duration (s)	1.124 ± 0.535	1.818 ± 0.789	−4.083	0.001 *
Saccade				
total saccade duration (s)	1.059 ± 0.176	1.597 ± 0.126	−20.207	0.001 *
saccade duration (s)	0.255 ± 0.104	0.314 ± 0.153	−3.734	0.001 *
Pupil				
pupil size	3.771 ± 0.023	3.767 ± 0.018	−0.743	0.984
AOI pupil size	3.895 ± 0.048	3.898 ± 0.043	−0.685	0.456
pupil difference	0.124 ± 0.045	0.131 ± 0.061	−0.423	0.338

* means $p < 0.05$.

5.2.2. Differences in Indoor Environments

In this section, we select gaze duration and saccade duration. The t-test results show that the participants pay significantly different amounts of attention to indoor landmark types between self-location and orientation (Figure 7).

The participants in the orientation group had a significantly longer gaze duration (mean = 2.242 s, SD = 0.547) and saccade duration (mean = 0.401 s, SD = 0.141) on store landmarks than did the participants in the self-location group (mean = 1.179 s, SD = 0.442; mean = 0.207 s, SD = 0.107). A similar trend also occurs for the elevator landmark. The mean gaze duration and saccade duration of the participants in the orientation group were 1.725 s (SD = 0.858) and 0.255 s (SD = 0.121), respectively, which were significantly higher than those in the self-location group, indicating that the store and elevator landmarks were significantly more attractive to the orientation group than to the self-location

group. Although the participants in the orientation group had a significantly shorter gaze duration on signs ($t = -2.698$, $p = 0.004$), the saccade duration on signs did not show a significant difference between the self-location and orientation groups ($t = -0.521$, $p = 0.309$).

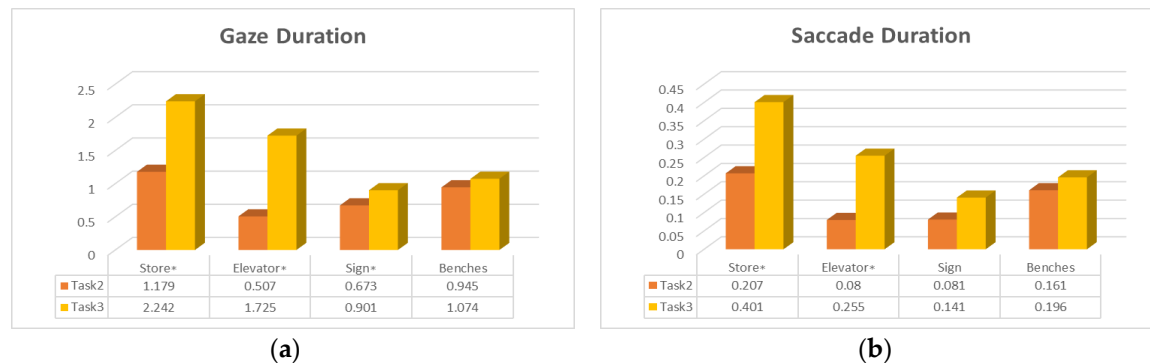


Figure 7. Comparison of the differences in visual salience among stores, elevators, signs and benches between tasks 2 and 3. * means $p < 0.05$. (a) Gaze duration differences between task 2 and 3; (b) Saccade duration differences between task 2 and 3.

There are two reasons for this phenomenon. First, we selected an indoor mall as the experimental environment; stores are the most important landmarks in a shopping centre, and participants are prone to observe store landmarks. Second, according to the landmark salience model, explicit marks play an important role in semantic salience, which indicates that it is easier for participants to find this factor attractive. Store landmarks have explicit marks, but the others do not.

5.3. Compared with Previous Research

Previous research has mainly calculated landmark salience by equal weighting, expert knowledge and instance-based scoring method. For example, the original salience model sets all weights as equal [7], and this method has been adapted by different salience models [17,26]. Nuhn [8] proposed a salience model that weights results using expert knowledge based on formal research. Zhu [19] constructed an instance-based scoring system to evaluate indoor landmark salience. Thus, the PLSR, equal weighting, expert knowledge and instance-based scoring methods were selected to compare the accuracy of landmark salience evaluation methods.

The landmark salience weighting results calculated by PLSR are mentioned in Section 4.2. The weighting results evaluated by the equal weighting method are shown in Table 11. As for the expert knowledge method, seven researchers with a PhD in cartography were invited to weight the importance of indoor landmark salience, and the results are shown in Table 11. Regarding the instance-based scoring method, the landmark salience attractiveness and weighting results are adapted from Zhu [19]. The landmark salience measured by the three algorithms was compared with the landmark salience based on eye-tracking data in Section 4.2.2. The difference value results are shown in Figure 8.

As presented in Figure 8, the circle size represents the landmark salience difference between eye-tracking data and four weighting methods. The larger the size of the circle, the bigger the difference value. The difference values of PLSR are lower than the other three weighting methods in both task 2 and 3. The highest difference value shown in AOI18 was calculated by equal weighting in both tasks 2 and 3. However, the difference value calculated by equal weighting is lower than those calculated by expert knowledge and instance-based scoring method in AOI13 in task 3. Thus, the most accurate weighting method in this study is PLSR. The accuracy of the other three weighting methods changes in different landmark attractiveness and tasks.

Table 11. Weighting results of four methods between tasks 2 and 3.

Visual	Shape	Deviation	Hue	Brightness	Façade area	Distance	Prominence	Unique label		
Task2										
PLSR	0.018	0.107	0.060	—	0.0003	−0.011	—	—		
Equal	0.166	0.166	0.166	0.166	0.160	0.166	—	—		
Expert	0.370	0.150	0.340	0.120	0.100	−0.080	—	—		
Instance	—	—	—	—	0.0565	—	0.3437	0.1394		
Task3										
PLSR	0.005	0.218	0.159	—	0.001	−0.014	—	—		
Equal	0.166	0.166	0.166	0.166	0.16	0.166	—	—		
Expert	0.350	0.210	0.280	0.230	0.090	−0.160	—	—		
Instance	—	—	—	—	0.0565	—	0.3437	0.1394		
Semantic	Importance	Marks	Familiarity	Uniqueness	Structural	Route	Object	Location	Spatial extent	Permanence
Task2					Task2					
PLSR	1.259	—	0.054	—	PLSR	0.048	—	0.228	—	—
Equal	0.334	0.333	0.333	—	Equal	0.334	0.333	0.333	—	—
Expert	0.290	0.380	0.330	—	Expert	0.280	0.300	0.420	—	—
Instance	—	0.0156	0.1070	0.0408	Instance	—	—	0.1988	0.0261	0.0721
Task3					Task3					
PLSR	3.362	—	−0.098	—	PLSR	0.088	—	0.339	—	—
Equal	0.334	0.333	0.333	—	Equal	0.334	0.333	0.333	—	—
Expert	0.230	0.350	0.420	—	Expert	0.250	0.280	0.470	—	—
Instance	—	0.0156	0.1070	0.0408	Instance	—	—	0.1988	0.0261	0.0721

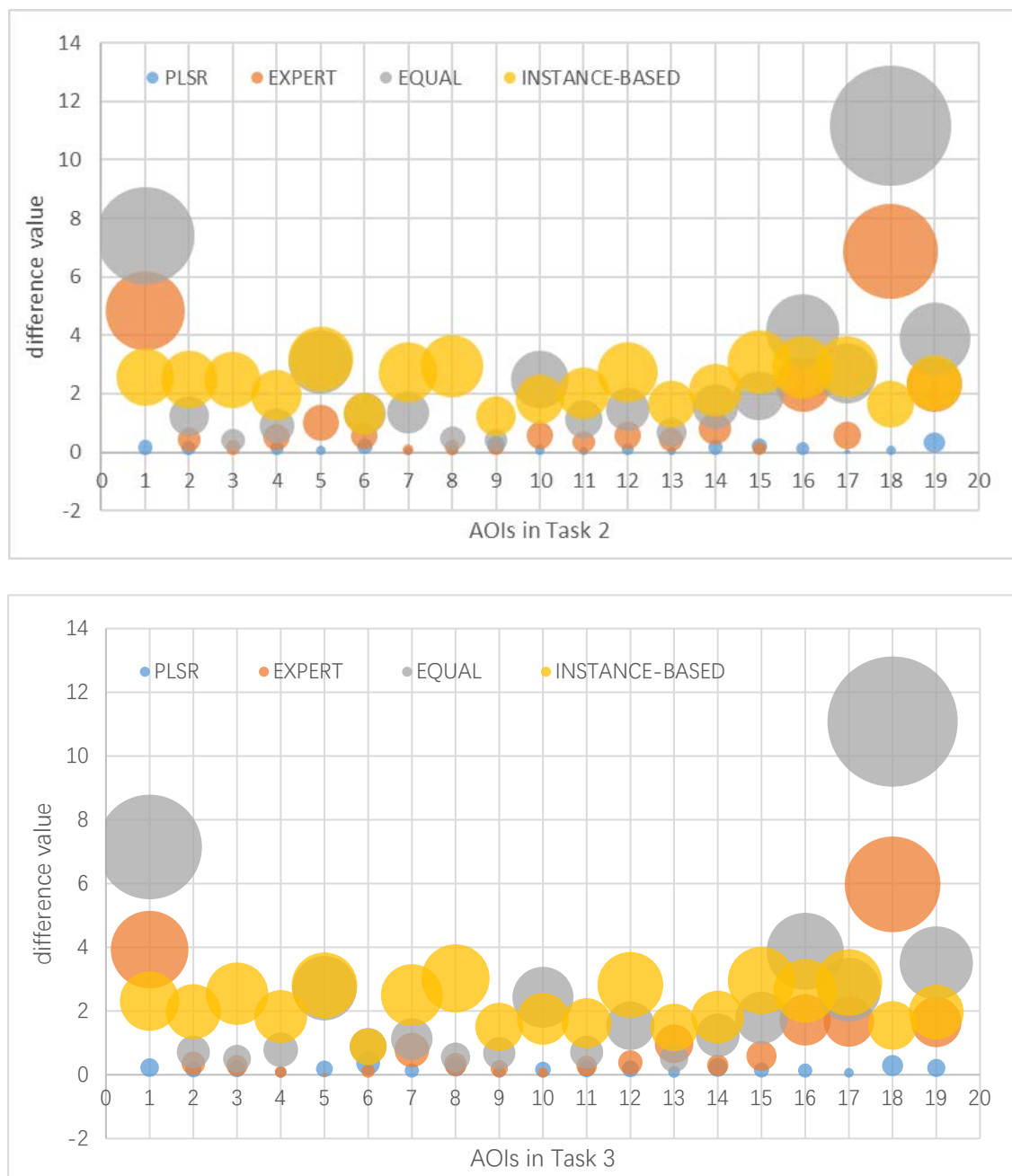


Figure 8. Comparison of four landmark salience calculation methods between task 2 and 3.

5.4. Improvements for Current Studies

Prior studies have considered landmark salience differences in personal dimensions, time dimensions (day and night) and environmental dimensions (indoor and outdoor), but few researchers have determined the differences in salience among various wayfinding tasks. For example, Nuhn and Timpf [8] defined the personal dimensions of landmarks, including spatial knowledge, interests, goals and backgrounds. Additionally, the salience of an object is different across people with various personal dimensions. Duckham, Winter, and Robinson [4] introduced nighttime vs. daytime factors into computing the salience of individual landmarks. Regarding the environmental dimension, researchers have proposed a salience model for both indoor and outdoor settings [24,27,28]. However, they could not determine the differences in salience in various tasks. This article provides a method to calculate

landmark salience based on participants' eye-tracking data, and the landmark salience models are different between self-location and orientation tasks.

The indoor landmark salience proposed in this article could be applied to design indoor maps for different tasks. For instance, we calculated landmark salience results in the first floor of the David Mall. Landmarks with high salience results could be regarded as areas of interest (AOI) [26]. Thus, we selected landmarks with salience results higher than the mean results as AOIs. The indoor maps for self-location and orientation were design by ArcMap 10.1 (Figure 9). Figure 9 shows that AOIs are different between indoor maps for self-location and orientation. For instance, Coach and BOTTEGA are AOIs in indoor maps for orientation, but they are not AOIs in indoor map for self-location.



Figure 9. Comparison of the indoor map between tasks 2 and 3. (a) Indoor map based on task 2; (b) Indoor map based on task 3.

In addition, the existing research has applied visual input technologies to realize human–computer interactions for navigation. Richter [17] pointed out that interaction between humans and computers has a specific focus on wayfinding. Additionally, landmark references have been proven to show the importance and benefits of this interaction. Landmark salience based on eye-tracking data can be theoretically inputted into real-time indoor navigation. For instance, participants can only passively receive landmark information when using traditional navigation applications. Thus, it is possible to construct a real-time gaze-aware navigation assistant that can actively detect the participant's eye movements [15], calculate the visual salience and recommend attractive landmarks to the participant in the future.

6. Conclusions and Future Research

This study aimed to establish an indoor landmark salience model based on eye-tracking data and to compare the differences in salience between self-location and orientation. The results show

two findings. Finding 1 proves that eye-tracking data could be used to measure indoor landmark salience. For instance, seven types of eye movement could be applied to analyse salience, and the salience result of the combined eye-tracking data was more accurate than that of the other types of eye movements. In addition, the PLSR weighting algorithm was more accurate than the other current weighting methods. Finding 2 shows that significant differences in landmark salience occurred between self-location and orientation. The participants paid more attention to landmarks that were stores and elevators in the orientation task. Thus, it is necessary to generate an indoor landmark salience model for different tasks. This study can contribute to the development of an indoor navigation map design in cartography and GIScience. For instance, landmarks with higher landmark salience could be highlighted in indoor navigation maps. In the meantime, it is meaningful for cartographers to redesign indoor maps for different wayfinding tasks.

However, since the experimental materials were statistical images, we did not discuss visual performance in real-world environments. The first challenge of indoor real-world experiments is how to calculate landmark salience attractiveness. For example, the information on shape factors, visual distance or façade areas keeps changing as the participants walk. It is difficult to define the exact visual factors of different landmarks. The second challenge is distraction caused by customers in a mall. Participants might be attracted by the people in real-world environments, and their concentration in the experiment may decrease.

As discussed in the previous section, future research can focus on the definition and calculation of landmark salience attractiveness in changing scenes. Moreover, future research can conduct experiments to detect the landmark salience attractiveness in various indoor scenes, such as airports, hospitals or conference centres.

Author Contributions: Formal analysis, Chengshun Wang and Yecheng Yuan; Funding acquisition, Yufen Chen; Methodology, Chengshun Wang; Visualization, Chengshun Wang and Shulei Zheng; Writing—original draft, Chengshun Wang; Writing—review & editing, Yufen Chen and Shuang Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the [National Natural Science Foundations of China] grant number [41701475, 41501507] and [The National High Technology Research and Development Program of China] grant number [2012AA12A404].

Acknowledgments: The authors would like to thank all the reviewers for their helpful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lynch, K. *The Image of the City*; M.I.T. Press: Cambridge, MA, USA, 1960; pp. 46–68.
2. Allen, G.L. Spatial abilities, cognitive maps, and wayfinding: Bases for individual differences in spatial cognition and behavior. *J. Wayfinding Behav.* **1999**, *9*, 46–80.
3. Schwering, A.; Krukar, J.; Li, R.; Anacta, V.J.; Fuest, S. Wayfinding through orientation. *Spat. Cogn. Comput.* **2017**, *17*, 273–303. [[CrossRef](#)]
4. Duckham, M.; Winter, S.; Robinson, M. Including landmarks in routing instructions. *J. Locat. Based Serv.* **2010**, *4*, 28–52. [[CrossRef](#)]
5. Piccardi, L.; Palmiero, M.; Bocchi, A.; Boccia, M.; Guariglia, C. How does environmental knowledge allow us to come back home? *J. Exp. Brain Res.* **2019**, *237*, 1811–1820. [[CrossRef](#)] [[PubMed](#)]
6. Albrecht, R.; von Stülpnagel, R. Memory for salient landmarks: Empirical findings and a cognitive model. In Proceedings of the 11th International Conference, Spatial Cognition 2018, Tübingen, Germany, 5–8 September 2018.
7. Raubal, M.; Winter, S. Enriching wayfinding instructions with local landmarks. In *Geographic Information Science*; Egenhofer, M.J., Mark, D.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2002.
8. Nuhn, E.; Timpf, S. A multidimensional model for selecting personalised landmarks. *J. Locat. Based Serv.* **2017**, *11*, 153–180. [[CrossRef](#)]
9. Anagnostopoulos, V.; Havlena, M.; Kiefer, P.; Giannopoulos, I.; Schindler, K.; Raubal, M. Gaze-Informed location-based services. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1770–1797. [[CrossRef](#)]

10. Erkan, İ. Examining wayfinding behaviours in architectural spaces using brain imaging with electroencephalography (EEG). *Archit. Sci. Rev.* **2018**, *61*, 410–428. [\[CrossRef\]](#)
11. Kiefer, P.; Giannopoulos, I.; Raubal, M. Where am I? Investigating map matching during selflocalization with mobile eye tracking in an urban environment. *J. Trans. GIS* **2014**, *18*, 660–686. [\[CrossRef\]](#)
12. Koletsis, E.; van Elzakker, C.P.; Kraak, M.J.; Cartwright, W.; Arrowsmith, C.; Field, K. An investigation into challenges experienced when route planning, navigating and wayfinding. *J. Int. J. Cartogr.* **2017**, *3*, 4–18. [\[CrossRef\]](#)
13. Brunyé, T.T.; Gardony, A.L.; Holmes, A.; Taylor, H.A. Spatial decision dynamics during wayfinding: Intersections prompt the decision-making process. *Cogn. Res. Princ. Implic.* **2018**, *3*, 13. [\[CrossRef\]](#)
14. Jia, F.; Tian, J.; Zhi, M. A visual Saliency model of landmark-based on virtual geographical experiments. *Acta Geod. Cartogr. Sin.* **2018**, *47*, 1114–1122. [\[CrossRef\]](#)
15. Liao, H.; Dong, W.; Huang, H.; Gartner, G.; Liu, H. Inferring user tasks in pedestrian navigation from eye movement data in real-world environments. *Int. J. Geogr. Inf. Sci.* **2018**, *33*, 739–763. [\[CrossRef\]](#)
16. Sorrows, M.E.; Hirtle, S.C. The nature of landmarks for real and electronic spaces. In *International Conference on Spatial Information Theory*; Springer: Berlin/Heidelberg, Germany, 1999.
17. Richter, K.F.; Winter, S. *Landmarks: GIScience for Intelligent Services*; Springer Publishing Company: New York, NY, USA, 2014.
18. Elias, B.; Paelke, V.; Kuhnt, S. Concepts for the cartographic visualization of landmarks. In *Location Based Services & Telecartography-Proceedings of the Symposium 2005, Geowissenschaftliche Mitteilungen*; Gartner, G., Ed.; Vienna University of Technology: Wien, Austria, 2005; pp. 1149–1155.
19. Zhu, L.; Svedová, H.; Shen, J.; Stachon, Z.; Shi, J.; Snopková, D.; Li, X. An instance-based scoring system for indoor landmark saliency evaluation. *Geografie* **2019**, *124*, 103–131. [\[CrossRef\]](#)
20. Mummidi, L.; Krumm, J. Discovering points of interest from users' map annotations. *GeoJournal* **2008**, *72*, 215–227. [\[CrossRef\]](#)
21. Wang, C.; Chen, Y.; Zheng, S.; Liao, H. Gender and Age Differences in Using Indoor Maps for Wayfinding in Real Environments. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 11. [\[CrossRef\]](#)
22. Huang, H.; Gartner, G.; Krisp, J.M.; Raubal, M.; Van de Nico, W. Location based services: Ongoing evolution and research agenda. *J. Locat. Based Serv.* **2018**, *12*, 63–93. [\[CrossRef\]](#)
23. Fellner, I.; Huang, H.; Gartner, G. Turn Left after the WC, and Use the Lift to Go to the 2nd Floor'—Generation of Landmark-Based Route Instructions for Indoor Navigation. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 183. [\[CrossRef\]](#)
24. Li, L.; Mao, K.; Li, G.; Wen, Y.A. A Landmark-based cognition strength grid model for indoor guidance. *Surv. Rev.* **2017**, *50*, 336–346. [\[CrossRef\]](#)
25. Gkonos, C.; Giannopoulos, I.; Raubal, M. Maps, vibration or gaze? Comparison of novel navigation assistance in indoor and outdoor environments. *J. Locat. Based Serv.* **2017**, *11*, 29–49. [\[CrossRef\]](#)
26. Ohm, C.; Müller, M.; Ludwig, B. Evaluating indoor pedestrian navigation interfaces using mobile eye tracking. *Spat. Cogn. Comput.* **2017**, *17*, 32. [\[CrossRef\]](#)
27. Lyu, H.; Yu, Z.; Meng, L. A Computational Method for Indoor Landmark Extraction. In *Progress in Location-Based Services 2014*; Springer: Cham, Germany, 2015; pp. 45–59.
28. GTze, J.; Boye, J. Learning landmark saliency models from users' route instructions. *J. Locat. Based Serv.* **2016**, *10*, 47–63. [\[CrossRef\]](#)
29. Fang, Z.; Li, Q.; Shaw, S.L. What about people in pedestrian navigation? *Geo Spat. Inf. Sci.* **2015**, *18*, 135–150. [\[CrossRef\]](#)
30. Golledge, R.G. Human wayfinding and cognitive maps. In *The Colonization of Unfamiliar Landscapes*; Routledge: London, UK, 1999; pp. 5–45.
31. Liao, H.; Dong, W. An exploratory study investigating gender effects on using 3d maps for spatial orientation in wayfinding. *J. Int. J. Geo Inf.* **2017**, *6*, 60. [\[CrossRef\]](#)
32. Meilinger, T.; Knauff, M. Ask for directions or use a map: A field experiment on spatial orientation and wayfinding in an urban environment. *J. Surv.* **2008**, *53*, 13–23. [\[CrossRef\]](#)
33. Wiener, J.M.; de Condappa, O.; Hölscher, C. Do you have to look where you go? Gaze behaviour during spatial decision making. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Boston, MA, USA, 20–23 July 2011; Cognitive Science Society: Austin, TX, USA, 2011.
34. Lscher, C.; Büchner, S.J.; Meilinger, T. Adaptivity of wayfinding strategies in a multi-building ensemble: The effects of spatial structure, task requirements, and metric information. *J. Environ. Psychol.* **2009**, *29*, 208–219.

35. Zhang, H.; Ye, C. An indoor wayfinding system based on geometric features aided graph SLAM for the visually impaired. *J. IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1592–1604. [[CrossRef](#)] [[PubMed](#)]
36. Kinsley, K.M.; Dan, S.; Spitler, J. GoPro as an ethnographic tool: A wayfinding study in an academic library. *J. Access Serv.* **2016**, *13*, 7–23. [[CrossRef](#)]
37. Tanriverdi, V.; Jacob, R.J.K. Interacting with eye movements in virtual environments. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, The Hague, The Netherlands, 1 April 2000; pp. 265–272.
38. Steinke, T. Eye movement studies in cartography and related fields. *J. Cartogr.* **1987**, *24*, 197–221. [[CrossRef](#)]
39. Ohm, C.; Müller, M.; Ludwig, B.; Bienk, S. Where is the landmark? Eye tracking studies in large-scale indoor environments. In Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research (in Conjunction with GIScience 2014), Vienna, Austria, 30 October 2014; Peter, K., Ioannis, G., Antonio, K., Raubal, M., Eds.; CEUR: Aachen, Germany, 2014; pp. 47–51.
40. Schrom-Feiertag, H.; Settgast, V.; Seer, S. Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment. *Spat. Cogn. Comput.* **2017**, *17*, 163–183. [[CrossRef](#)]
41. Kiefer, P.; Giannopoulos, I.; Raubal, M.; Duchowski, A.T. Eye tracking for spatial research: Cognition, computation, challenges. *Spat. Cogn. Comput.* **2017**, *17*, 1–9. [[CrossRef](#)]
42. Bulling, A.; Ward, J.A.; Gellersen, H.; Troster, G. Eye movement analysis for activity recognition using electrooculography. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 741–753. [[CrossRef](#)] [[PubMed](#)]
43. Kiefer, P.; Giannopoulos, I.; Raubal, M. Using eye movements to recognize activities on cartographic maps. In Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Orlando, FL, USA, 5–8 November 2013; pp. 478–481.
44. Goldberg, J.H.; Kotval, X.P. Computer interface evaluation using eye movements: Methods and constructs. *Int. J. Ind. Ergon.* **1999**, *24*, 631–645. [[CrossRef](#)]
45. Dong, W.; Qin, T.; Liao, H.; Liu, Y.; Liu, J. Comparing the roles of landmark visual salience and semantic salience in visual guidance during indoor wayfinding. *Cartogr. Geogr. Inf. Sci.* **2019**, 1–15. [[CrossRef](#)]
46. Li, W.; Liu, Z. A method of SVM with normalization in intrusion detection. *Procedia Environ. Sci.* **2011**, *11*, 256–262. [[CrossRef](#)]
47. Just, M.A.; Carpenter, P.A. Eye fixations and cognitive processes. *Cogn. Psychol.* **1976**, *8*, 441–480. [[CrossRef](#)]
48. Dong, W.; Liao, H.; Roth, R.E.; Wang, S. Eye tracking to explore the potential of enhanced imagery basemaps in web mapping. *Cartogr. J.* **2014**, *51*, 313–329. [[CrossRef](#)]
49. Nuhn, E.; Timpf, S. Personal dimensions of landmarks. In Proceedings of the International Conference on Geographic Information Science, Wageningen, The Netherlands, 9–12 May 2017.
50. Liben, L.; Myers, L.; Christensen, A. Identifying locations and directions on field and representational mapping tasks: Predictors of success. *Spat. Cogn. Comput.* **2010**, *10*, 105–134. [[CrossRef](#)]
51. Liao, H.; Dong, W.; Peng, C.; Liu, H. Exploring differences of visual attention in pedestrian navigation when using 2D maps and 3D geo-browsers. *Cartogr. Geogr. Inf. Sci.* **2016**, *44*, 474–490. [[CrossRef](#)]

