VISUAL SEMANTIC PARAMETERIZATION – TO ENHANCE BLIND USER PERCEPTION FOR INDOOR NAVIGATION

Samleo L . Joseph, Chucai Yi, Jizhong Xiao, YingLi Tian[*]

Department of Electrical Engineering
The City University of New York - City College
Convent Ave at 140th Street, New York, USA.
(sjoseph, jxiao, ytian)@ccny.cuny.edu; cyi@gc.cuy.edu

Fei Yan [†]

Research Center of Information
and Control, Dalian university
of Technology, Dalian, China.
catyanfei@gmail.com

ABSTRACT

This paper presents a novel approach of utilizing the floor plan maps posted on the buildings to infer a semantic plan that aids in the navigation of a visually impaired person. The extracted landmarks such as room numbers, doors, etc act as a parameter to infer the way points to each room. This provides a mental mapping of the environment to design a navigation framework for future use. A human motion model is used to predict a path based on how real humans ambulate towards a goal by avoiding obstacles. Travel route is presented in terms of blind understandable units, which is achieved by accurate estimation of the user's location and confirmed by extracting the landmarks posted on the doors. The results show that it is feasible to make a blind user to travel independently by providing the constraints required for safe navigation.

***Index Terms*—** floor plan, semantic schema, human motion, signage, navigation

1. INTRODUCTION

Millions of people live in this world with incapacities of understanding the environment due to visual impairment. Individuals with normal vision, who view a floor-plan in order to navigate to a room of interest, can infer a shortest route to safely reach a destination and confirm the travelled route based on the identification of landmarks. Whereas, in case of individuals with visual impairment, its highly challenging to make them independently navigate to a destination. So, research in the field of navigation, applied to power the mobility of visually challenged has concerned sophisticated technology and techniques. A more effective approach is to solve this problem in real-time, by taking advantage of semantic cues in the surrounding environment.

In robotics, a precise model of the relationship between control parameters and body kinematics is designed as a motion model. This is influenced by physical constraints in the environment in order to navigate a robot to next step towards the destination. In contrast, a human acquires control parameters that yield safe navigation from the physical constraints of the environment. If these constraints change, a Human may adapt to optimize those parameters until path is stabilized again. Whereas, a visually impaired person is still a human, who can actively adapt to the environmental parameters, provided these constraints are precisely nurtured, in a blind perceivable manner. So, this paper focus on nurturing those constraints required for safe navigation of a visually impaired person.

The challenges involved in designing a navigation framework utilizing a floor plan are as follows:

- How to design a schema that nurtures the constraint of an environment required for a blind navigation?
- How to emphasize the semantic cues inferred from a floor plan to design a meaningful schema?
- Can the proposed schema be created using analogous floor-plans of all the buildings in the world?
- How to present the travel distance in terms of blind understandable units to notify the intended travel route?
- How to predict a human walk and plan a path by avoiding the obstacles on the way towards the destination?

This paper solves the above mentioned challenges by modeling the constrains influencing in the environment for the safe navigation of the blind people.

2. LITERATURE REVIEW

In order to provide assistance for the blind navigation similar to a normal visioned person (who can visualize clues inherent in the building and navigate), several systems are developed using the conventional approach of generating a travel route by utilizing a precise 'pre-build map' maintained in its

---

[*] This work is supported in part by U.S. National Science Foundation (grants No. IIS- 0644127 and CBET-1160046), Federal HighWay Administration (DTFH61-12-H-00002) and PSC-CUNY (65789-00-43).

[†] The fifth author performed the work while at the City College.

database. This interaction style of manually feeding the constraints in the environment and then initiating the navigation process is referred as mixed-initiative modeling of navigation.

The system providers need to manually create a building infrastructure and feed the floor-plan into the system after explicitly annotating the landmarks within the map, before the blind user initiates a navigation task. Apostolopoulos and Fallah et al proposed Navatar that requires collection of building map and manual preprocessing to annotate landmarks required for navigation [1, 2]. The advantage of this mixed-initiative model is to utilize a powerful desktop computer that can augment a map with meaningful information within a short timespan.

Lee et al proposed universal circulation network for the wheel-chair access in architecture planning. It is derived from the Building information model (BIM) by using the door points to generate a graph plan [3]. Karimi et al proposed an approach in universal navigation on smartphones. It requires an interpolation scheme that uses spatial data on the end nodes of hallway segments. Geocodes that are computed from the co-ordinates of Point-of-interests are retrieved from a navigation database [4]. Lorenz et al proposed a navigation model with qualitative and quantitative labels for representing nodes (rooms and corridors) and edges (doors and passways). It uses access points to elevators to reach different floors from those interface nodes, provided by a graph hierarchy algorithm [5]. The main drawback with all these approaches are that the building information needs to be collected from the building owners and its meaningful information needs to be manually integrated into their system in order to generate a metric map for the indoor navigation.

In robotic literature, a conventional motion model called 'constant velocity model' is employed to predict next step towards destination, depending on robot's kinematics and control parameters influenced by the environmental constrains. Apostolopoulos and Fallah et al proposed Navatar that employes a constant velocity model with a 'minimal distance of a human walk metric' without considering the obstacles on the way in the prediction step [1, 2]. These motion models are prone to be inaccurate if the real human diverges from this model because of an interaction with obstacle on the surrounding environment.

The idea of solving the limitations and incapacities of the human misunderstanding due to the visual impairment is still a open problem. We propose a novel approach of indoor navigation by nurturing the environmental constrains that are inferenced from the visual cues, and planning a realistic path provided by a human motion model specifically for the ambulation of a blind person.

3. PROBLEM FORMULATION

Building a map from an unknown environment is a challenging problem and often requires complex task of acquiring the surroundings with sensors that approximately senses the environment in order to build a map. In the literature, research in the field of probabilistic robotics have solved this mapping approach using the simultaneous localization and mapping (SLAM) algorithm from the data acquired by sensors traversing along the entire environment [6]. But in our approach, we design a navigation framework called 'semantic schema' by inferring the meaningful information encrypted in the snapshot of a floor-map and moreover, the designed schema will be updated or corrected when the human-user traverses to the corresponding locations on that map.

4. SYSTEM OVERVIEW

The process chart of the proposed framework to empower the blind user perception for indoor navigation is shown in Figure 1. The system triggers the head-mounted smart glass which is comprised of camera and inertial measurement unit (IMU). If a normally sighted person identifies a floor plan posted on the building, then a snapshot of a floor plan can be acquired using the camera and provided to the blind user. The floor plan parametrization (section 5) acts as a visual perception unit for a blind user. It employs a heuristic method of extracting room numbers and door shapes from a raw floor plan data, which further acts as a parameter for defining an entry point to each room. Semantic schemata (section 6) acts as a central processing unit of our proposed system. It perceives a mental mapping of the environment and organizes the semantic information to generate a navigation framework called 'semantic schema' for the future use. When the user specifies a room of interest, both the qualitative and the quantitative information of the desired room is inferred from the semantic schema. It also provides the shortest route to reach a destination, including all the landmarks within the intended route. Human motion parametrization (section 7) solves the local obstacle avoidance problem by predicting a human path to reach those landmarks using a real-time path planning in local map. Pathing map parametrization (section 8) acts as a guidance module to inform the travel distance information in terms of blind understandable metric units which is achieved by accurately estimating the current position of the user. It integrates algorithms from the robotic SLAM, which uses the state prediction from the Human motion model and corrects the current location using the measurement updates from the landmarks. Finally, if the user reached a desired destination, verbal notification is activated to inform the user location.

5. FLOOR PLAN PARAMETERIZATION

To extract layout information from a floor map, we should localize the rooms and hall paths. However, it is difficult to extract room boundaries accurately, because of their complex intersections and merges. Thus we design a heuristic method,

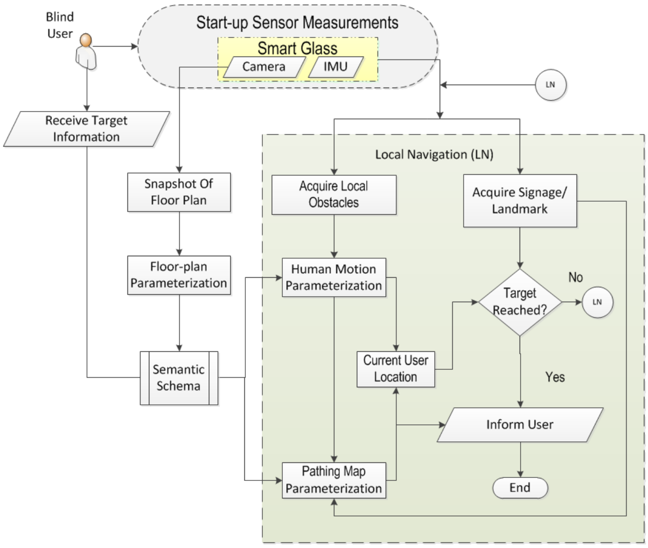


Fig. 1. Process chart of the proposed approach

which employs room numbers and door shapes to figure out the floor plan.

This method can be divided into two steps, room number detection and door shape detection. As shown in Figure 2(a) a room appears in the form of outbound rectangle and room number label. The label consists of three digital numbers in horizontal alignment. Thus we propose a rule-based method to localize the positions of all room number labels. First, canny edge detection is applied to obtain the edge map, as shown in Figure 2(b). Second, the boundaries that are composed of connected edge pixels are extracted from the edge map. Each boundary is assigned a bounding box with compatible size. Third, for each bounding box, we check whether it is located at the middle of two neighboring bounding boxes in similar height and horizontal alignment. If yes, we merge the three bounding boxes into a boundary group [1] (see Figure 3). Fourth, each boundary group is extended into a room number region, as shown in Figure 4).

Based on the detected regions of room numbers, we search for the range of the rooms and positions of the room doors in the floor map. As our observation, the door in the floor map is mostly in the form of D-shape, as shown in Figure 5. To detect the D-shape, a horizontal and a vertical scan are respectively generated from the region of room number. If the scan line touches the D-shape, one of its ends will have a monotonous variation. Therefore, we could detect a rough position of the D shape door, as shown in Figure 5. Assuming that all the doors correspond to the hall path of this floor, we generate anchor point by using the room number label and the D shape door.

(a) Snapshot of floor plan
(b) Edge map

Fig. 2. The floor map and the corresponding edge map.

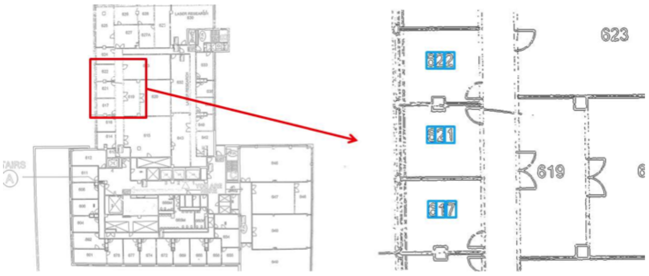


Fig. 3. The boundary group obtained from edge map.

6. SEMANTIC SCHEMATA

In order to provide a navigation framework for the future understanding of the environment, we use schemata to organize the knowledge acquired on the floor plans from section 5. So, semantic schemata acts as a central processing unit of our proposed system which is used to perceive a mental mapping of the environment.

The extracted meaningful information such as room number labels and its entry points on the hallway boundaries obtained from section 5, are used to generate a graph of nodes. A precise plan or schema is designed using those nodes which consists of all the waypoints to enter a room of interest in the building. Thus, a database with both the quantitative and the qualitative information is maintained to provide knowledge on the following: (1) identify a room of interest, (2) its location on the global map, and (3) all the neighboring waypoints connected to it. These information can be used to explore the shortest path using conventional path planning techniques.

7. HUMAN MOTION PARAMETERIZATION

In order to navigate to a desired room, a blind person might need to pass through other rooms on the intended travel route provided by the semantic schema. These passing room numbers can act as a landmark to the blind person and confirm the user's planned route. So, the intended path towards each door

Fig. 4. The detected regions of room numbers marked in red (best viewed in color version).

is emphasized as a short term goal. The problem of reaching each short goals inferred from the semantic schema can be solved by a real-time path planning in local map (until the long term goal is reached) which is considered as an obstacle avoidance problem.

To reach each short goals towards the nearest landmarks, a bio-inspired motion model is required to predict the user path. This motion model should effectively repel from the obstacles and safely navigate towards the nearest landmark. Hence, a human motion model originally proposed by Fajen et al. [7] is employed to imitate how humans walk towards the goal by avoiding obstacles using behavioral dynamics.

An illustration of the proposed working scenario with the model parameters involved in the blind inferenced motion model is shown in Figure 6. Here, we consider a scenario where a blind person intends to start from the exit in order to navigate towards the nearest landmark (restroom). A typical floor plan view with a blind person on the intended travel path in a local framework, is augmented with real-time obstacles acquired by a wearable head mounted camera. This is integrated into the semantic schema to predict the state of a Human model inferred from the following parametrization:

- the heading direction, Θ of the blind person (blue dotted line) with respect to the reference frame (black segmented line)
- the orientation, Ψ_g and distance, d_g towards the goal (green dotted line) with respect to the reference frame;

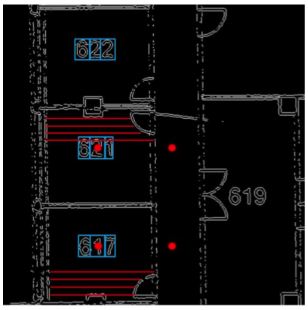


Fig. 5. D-Shape detection is performed to localize the door position. According to the room number region and door position, we could calculate the point in hall path.

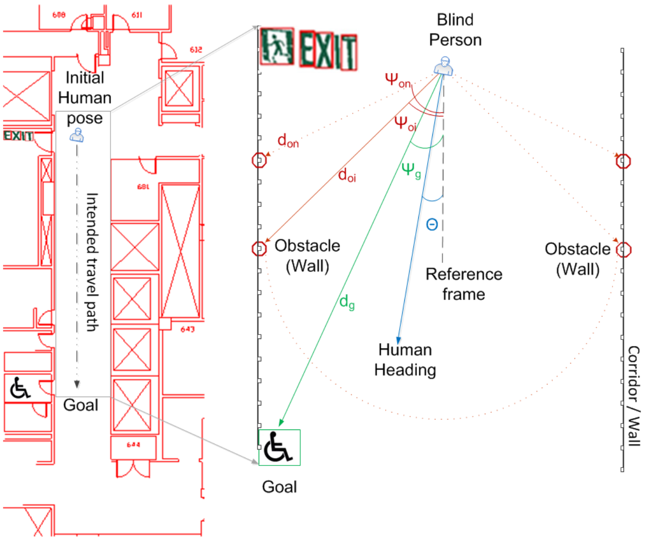


Fig. 6. Illustration of the proposed working scenario with the model parameters involved in the blind inferenced motion model.

- the orientation, Ψ_{o_i} and distance, d_{o_i} towards the obstacles (red dotted lines) with respect to the reference frame, where i is the number of obstacles;
- intended signage of the restroom to confirm the short term goal.

Thus, the real-time obstacle avoidance problem is solved by planning a path in local map using the following human motion model,

$$\frac{\partial^2\Theta}{\partial t^2} = -f_d\left(\frac{\partial\Theta}{\partial t}\right) - f_g(\Theta - \Psi_g, d_g) + \sum_{i=1}^{n} f_o(\Theta - \Psi_{oi}, d_{oi})$$

Where f_d, f_g, f_o are damping, goal and obstacle components, respectively. So, Ψ_g, d_g, Ψ_{o_i} and d_{o_i} change as the position of the human changes and a new path is generated.

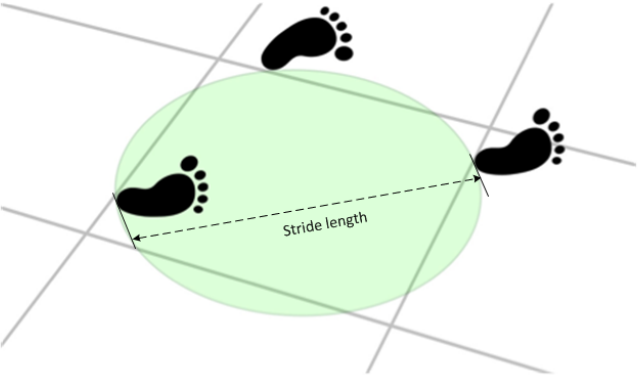


Fig. 7. A typical walking cycle in pathing unit.

8. PATHING MAP PARAMETERIZATION

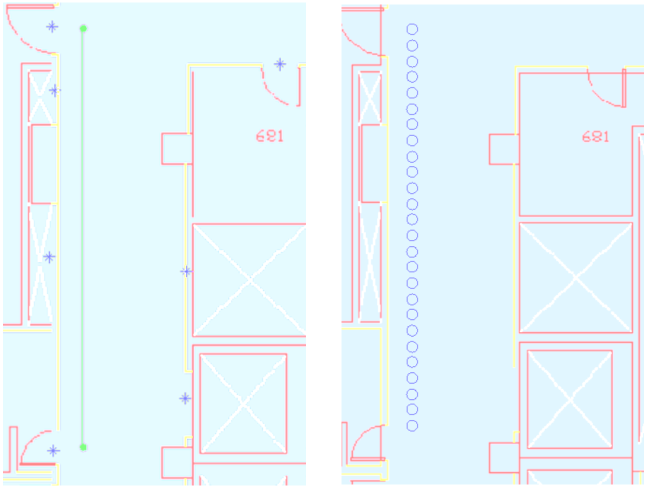
Waking length of a human varies for each individuals, so, an individual parameter is required to notify the remaining steps to be taken from the current location to a desired destination. Pathing map parameterization acts as a guidance module to present the travel distance information in terms of blind understandable metric units referred as pathing unit.

A pathing unit is defined as a circle with a diameter equal to the stride length of the blind user based on the distance and gait parameters of a typical walking cycle, as shown in Figure 7. This stride length is easily estimated by an individuals height times k with some negligible displacement error in centimeters, where k is 0.415 for men and 0.413 for women [8]. A pathing map is represented by the allocation of a track that is laid down for the ambulation of the blind person in a local map calibrated in pathing units.

In order to notify the travel distance information, an accurate estimation of the current user location is required. This can be achieved by using conventional robotic SLAM algorithms which updates and corrects the semantic schema when the user traverses to the corresponding locations on the global map. The location of the user, predicted using the human motion model can be used in the prediction step. The measurement updates from the landmarks can be used to correct the current location of the user. The semantic information acquired from the landmarks such as room numbers or signage can be used to provide audio feedback by informing their current location while the user passes the landmarks on the way towards the goal.

9. RESULTS AND DISCUSSION

A real floor plan posted near our lab is snapshot and processed to infer the meaningful information encrypted on it using our proposed approach of visual sematic parameterization. Based on the inferenced schema, a human motion is simulated on the floor map to reveal the fact that how a real blind person can navigate independently and safely reach a destination after understanding the surrounding environment.



(a) Intended travel route (b) Proposed pathing map

Fig. 8. Intended travel route designed by semantic schema.

Consider a navigation task as illustrated in Figure 6, where a blind person intends to start from the exit in order to navigate towards the nearest landmark (restroom). This task of navigating to a short term goal using the proposed navigation approach is as follows:

When a goal is fed into the semantic schemata, it designs a schema with the linked landmarks and the shortest route to navigate from the exit to reach the restroom. A typical floor plan view with an intended travel path is shown in Figure 8 (a), where the green line indicates the shortest path planned by the semantic schema.

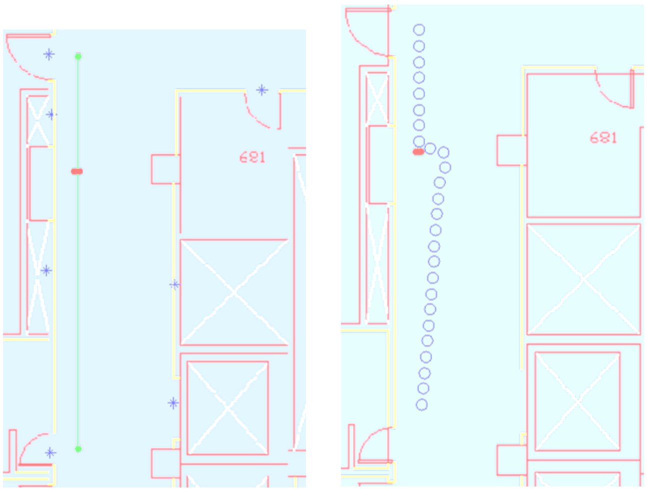
The current location of the blind user is inferred based on the visual cue, shown in Figure 9 (a). The location of the user is confirmed after extracting the encrypted location cue as ‘exit’, see Figure 9 (b). (For extraction algorithm, please refer the authors previous work on [9]) Now, the head-mounted smart glass is used to further capture the scenario, viewed by the user to identify any obstacles as considered in following two cases.

In case (1), consider a travel path, where there is no local obstacles other that the corridor along the intended route, planned by the schema. So, the human motion model predicts a straight path along the corridor. Then, the pathing map parametrization presents a travel route in pathing units, as shown in Figure 8 (b).

In case (2), consider a travel path, where there is a local obstacle identified along the intended route previously planned by the schema. Here, human motion model repels the obstacles and predicts a path around the obstacle on the way towards the destination. Then, the pathing map parametrization alters the travel route based on the physical constrains in the environment, as shown in Figure 10 (b).

(a) Snapshot of scene (b) Extracted visual cues

Fig. 9. Snapshot of initial and final scene with inferred cues.



(a) Intended travel route (b) Optimized Pathing map

Fig. 10. Intended travel route inferring obstacle on the way.

Finally, if the user reaches the destination, the final location is confirmed after inferring the visual cue, as shown in Figure 9. The location of the user is confirmed after extracting the encrypted location cue as ‘Women’, see Figure 9 (b).

The same process of navigating to a short goal, can be augmented with new neighboring landmarks (new short goals) to reach a long term goal based on the intended travel route designed by the semantic schema. This also solves the kidnap problem, where if the blind user gets lost, the system will be able to recognize a previously visited place and identify the current location to further plan a different route that reaches the desired goal position.

10. CONCLUSION AND FUTURE TRENDS

Our proposed approach revealed the fact that it is possible to make a blind person to navigate independently and safely in order to reach a destination, provided – the person is augmented with our proposed navigation framework to enhance understanding of the surrounding environment. Moreover, when the blind user gets lost, the system will be able to recognize a previously visited location and further plan a different route to reach the desired goal position. This can promote independent living of the blind person.

Our future work is to test the proposed approach with a blind subject after integrating sensors such as Kinect (for accurate obstacle detection), speech module, tactile feedback, etc. which lead to a completely functional navigation device to assist the blind user. We will also integrate crowd-sensing effects to provide awareness to the blind user based on the evidence of some interesting pattern inferred using the crowd analytics.

11. REFERENCES

[1] Ilias Apostolopoulos, Navid Fallah, Eelke Folmer, and Kostas E. Bekris, “Integrated online localization and navigation for people with visual impairments using smart phones,” in *International Conference on Robotics and Automation (ICRA)*, 2012.

[2] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer, “The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks,” in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, 2012.

[3] Jin-kook Lee, Charles M Eastman, Jaemin Lee, Matti Kannala, and Yeon-suk Jeong, “Computing walking distances within buildings using the universal circulation network,” *Environment and Planning B: Planning and Design*, vol. 37, no. 4, 2010.

[4] Hassan A. Karimi, *Universal Navigation on Smartphones*, Springer, 2011.

[5] Bernhard Lorenz, Hans Ohlbach, and Edgar-Philipp Stoffel, “A hybrid spatial model for representing indoor environments,” *Web and Wireless Geographical Information Systems*, 2006.

[6] Sebastian Thrun, Wolfram Burgard, and Dieter Fox, *Probabilistic Robotics*, The MIT Press, Aug. 2005.

[7] Brett R. Fajen, William H. Warren, Selim Temizer, and Leslie Pack Kaelbling, “A dynamical model of visually-guided steering, obstacle avoidance, and route selection,” *International Journal of Computer Vision*, vol. 54, no. 1-3, 2003.

[8] A.R. Pratama, Widyawan, and R. Hidayat, “Smartphone-based pedestrian dead reckoning as an indoor positioning system,” in *2012 International Conference on System Engineering and Technology (ICSET)*, 2012.

[9] YingLi Tian, Xiaodong Yang, Chucai Yi, and Aries Arditi, “Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments,” *Machine Vision and Applications*, pp. 1–15, 2012.