Poster: An Infrastructureless and Self-deployable Indoor Navigation Approach Using Semantic Signatures

Taeyu Im
Stony Brook University
quocduy.vo@stonybrook.edu

Pradipta De
Georgia Southern University
pde@georgiasouthern.edu

ABSTRACT

Semantic localization refers to the process of finding one's location with respect to visible or identifiable objects in the scene instead of finding position coordinates. In this research, we use textual signs and their geographical relationships inside a building as semantic signatures to design an infrastructureless indoor navigation system without a site survey. We propose a computer vision-based approach, which takes as an input a floor plan image and automatically infers the floor graph. The constructed graph is then used to estimate the shortest path, which is a sequence of store names, that the user needs to pass to reach her destination.

Keywords

Semantic Signature, Indoor Navigation, Android Application, Computer Vision

1. INTRODUCTION

Fingerprint-based indoor navigation using smartphones has become important in the past decade [1]. A typical system consists of four main components: environment representation (ER), localization, path planning, and user interaction. Among them, ER is the most important component which stores and retrieves different types of information in the user's surroundings. Depending on the ER approach, the stored information can be differently used in localization, path planing, and providing navigation instructions.

Although advanced techniques for environment representation have been developed, various shortcomings still remain. Most ER methods used in existing navigation systems require a process of site survey where fingerprints at different reference points are collected. However, this process is time consuming, labor intensive, and easily affected by environmental dynamics. Although, crowdsourcing has been used to reduce the costly site survey with affordable prices, it is slow to adapt to changes in the environment, requires a survey training for the users, and also depends on the willingness of the users to join [2]. The deployment of such systems in large-scale therefore remains a challenge.

MobiCom'16 October 03-07, 2016, New York City, NY, USA

ACM ISBN 978-1-4503-4226-1/16/10.

DOI: http://dx.doi.org/10.1145/2973750.2985270

Also, most indoor localization methods require physical infrastructures (e.g., Wi-Fi access points, or markers). Path planning methods focus on estimating the shortest path or shortest travel time. However, tourists often prefer longer paths that have points of interest en route. Lastly, existing navigation techniques use many sensors to continuously sense the environment or a display camera to interact with the system. This rapidly drains the battery.

Recently, objects inside a building and their spatial relationships have been used as semantic signatures for navigation [3]. Detecting such objects requires high computational cost from the devices, and leads to rapidly draining the battery. In this paper, we exploit the use of textual signs (e.g., store names, AIGA signs) and their spatial relationships as semantic signatures for navigation. For instance, a sequence of store names (`Starbucks`, `KFC`, etc.) along hallways is invariant and can be used to construct a route for navigation.

Algorithm 1 *Graph Construction*

Require: A captured image of a floor plan I

for each pixel $p \in I$ **do**
    p = *ImageSmoothing* (p, I) /* Image Pre-processing */
end for
Set *blobs* = *BlobsDetection*(I)
hallway = *HallwayDetection*(I) /* Detecting hallway */
for each blob $b \in$ *blobs* **do**
    roomName rn = *OCR* (b) /* Detecting room names */
    TextsRemoval(b) // Removing all texts in the image
end for
Set corners = *CornersDetection* (I)
Set *selectedCorners* = [ ] /*Selecting corners on the hallway*/
for each corner $c \in$ *corners* **do**
    if IsOnHallway(c) **then**
        selectedCorners.put(c)
    end if
end for
refPoints = *MidpointsCompute*(*selectedCorners*, *hallway*)
graph = *GraphConstruction*(*refPoints*, *hallway*)

Since floor plans are widely available in public places like shopping malls or airports, we propose a computer vision-based technique which explores the ontologies of textual signs. It takes as an input a floor plan image captured by smartphone camera and returns a graph which represents the relationships of stores inside the building. Using this graph, we plan the path to calculate the navigation route, which is a sequence of store names. The user needs to sequentially pass all the store in the list starting from her current location until she arrives at her destination. The current location can be computed while interpreting the floor plan, or it can be found using store names.

2. PROPOSED SOLUTION

The proposed system consists of two components: envi-

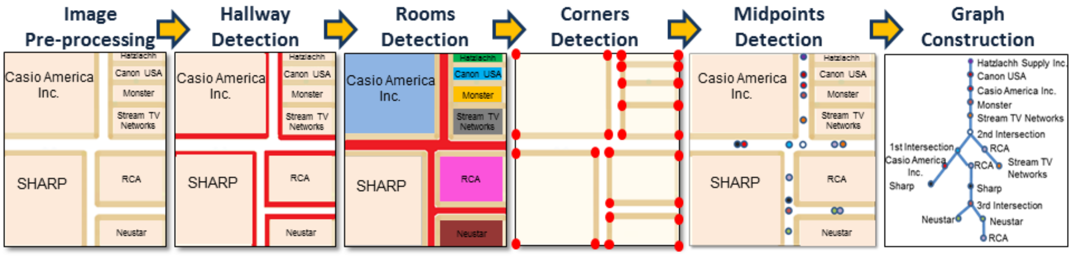

Figure 1: The workflow of graph construction process with an example

ronment representation (ER) and navigation. The ER component constructs a weighted graph from a floor plan image captured by the user at one of entrances to the building. The navigation component uses the constructed graph to locate the user, plan the path, and provide navigation instructions.

2.1 Environment Representation (ER)

To construct the graph, we first extract structures of the building from the floor plan such as hallways, rooms, and intersections. As shown in Figure 1, we first use image smoothing techniques to make the boundary and individual rooms clear. Next, we find all blobs in the floor image using OpenCV. The blobs can be either individual stores or hallways. Hallways can be recognized based on their sizes relative to the whole image and rooms. After detecting the hallways, we detect rooms located on each hallway by checking corresponding blobs along each hallway. Since each room has its name located inside the corresponding blob, we apply optical character recognition on its blob to detect its name.

Algorithm 2 *Semantic Navigation*

Require: *route*=$\{P_0, P_1, ..., P_n\}$ with turns are marked.
$P_i = P_0$
while (P_i is not the last turn P_k) {
  showDirection(P_i) //shows which way the user have to take
  move(P_i,P_{i+1}) //user selects the way and to the next point
  if (*isCorrectWay*(P_{i+1}, *side*)) { // user took the right way
    P_i = *nextTurn*(P_{i+1}) //user walks to reach the next turn
  } **else** { //User took the wrong way
    P_0=*localization*() //identify the user's current location
    route=*computeRoute*(P_0, P_n) //re-compute the route
    navigation(route) // start navigation again from P_{i+1}
  }
move(P_i,P_n) //move to the last point

We next detect corners in the image. Due to the existence of texts, the corners detection process needs more computation. Thus, we remove all of the texts appearing in the blobs before detecting corners. However, the corner detection process returns all of the corners appearing in the image including corners which are not located on any hallway. Since our target is to find reference points (RP), which represent entrances to stores on the map, corners which are not located on any hallway will be ignored. Finally, a RP is measured by computing the midpoint between two adjacent corners located on the same hallway.

Each store has different number of RPs. To represent the relative locations between semantic fingerprints, we model a physical floor graph with an weighted graph $G_{PF} = (V, E)$, where each vertex $v \in V$ is a RP (or a point of interest), each edge $(u, v) \in E$ indicates the reachability (or direct connection) between two RPs u and v, and a weight of each edge is the relative distance (in pixels) between two vertexes.

2.2 Navigation

While interpreting the floor plan, we also locate the current location of the user by detecting a red dot on the map indicating the user's current location. Given the current location and the destination inputted by the user, the system first plans for the shortest path using Dijkstra's shortest path algorithm. The path planning returns a sequence of RPs $(P_0, P_1, ..., P_n)$, where P_0 is the user's current location and P_n is the user-specified destination.

At the starting point (P_0), as shown in Algorithm 2, the current location is detected and presented to the user together with the direction. When the user arrives at an intersection, however, she may get confused to select the correct way because the store names might be hidden from her view. We address this problem by marking intersections as turning points while constructing floor graph. Starting from each intersection, we also mark the side of stores located next to the intersections. Whenever the user arrives at an intersection (P_i), the system shows the current intersection and also the way to take. However, the user may take a wrong way. She needs to confirm if she took the right way or not by checking the first store (P_{i+1}) she sees and the side it is located (on her left-hand side or right-hand side). If the user took the wrong way, she is required to take a picture of store names around her for localization. Using the store names extracted from the captured picture, the system can identify the user's current location and re-compute the route for navigation.

3. CONCLUSION

We propose an indoor navigation system using semantic signature without a site survey or a calibration process. We construct a floor graph from a floor plan image by exploiting the spatial relationships of rooms inside a building. Using the floor graph, we estimate a sequence of store names as the shortest path to navigate the user.

Acknowledgement

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the "ICT Consilience Creative Program" (IITP-2015-R0346-15-1007) supervised by the IITP (Institute for Information and Communications Technology Promotion)"

4. REFERENCES

[1] N. Fallah, I. Apostolopoulos, K. Bekris, and E. Folmer. Indoor human navigation systems: A survey. *Interacting with Computers*, 2013.

[2] R. Gao, M. Zhao, T. Ye, F. Ye, Y. Wang, K. Bian, T. Wang, and X. Li. Jigsaw: Indoor floor plan reconstruction via mobile crowdsensing. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom '14, New York,USA, 2014. ACM.

[3] I. Kostavelis and A. Gasteratos. Semantic mapping for mobile robotics tasks: A survey. *Robotics and Autonomous Systems*, 66:86 – 103, 2015.