Demo: Assisting Visually Impaired People Navigate Indoors*

J. Pablo Muñoz
City University of New York, NY
jmunoz2@gradcenter.cuny.edu

Bing Li, Xuejian Rong
Jizhong Xiao,
Yingli Tian
City College of New York, NY
{bli, jxiao, xrong, ytian}@ccny.cuny.edu

Aries Arditi
Visibility Metrics
Chappaqua, NY
arditi@visibilitymetrics.com

Abstract

Research in Artificial Intelligence, Robotics and Computer Vision has recently made great strides in improving indoor localization. Publicly available technology now allows for indoor localization with very small margins of error. In this demo, we show a system that uses state-of the-art technology to assist visually impaired people navigate indoors. Our system takes advantage of spatial representations from CAD files, or floor plan images, to extract valuable information that later can be used to improve navigation and human-computer interaction. Using depth information, our system is capable of detecting obstacles and guiding the user to avoid them.

1 Introduction

Unknown indoor environments present a great difficulty for visually impaired people. Current sensor technology [Lee and Dugan, 2016] can be used to assist this segment of the population navigate indoors. We demonstrate a system that is capable of guiding visually impaired users to reach destinations in multi-floor environments. Our system incorporates well-known Artificial Intelligence algorithms, e.g., A*, and modifies their output to ensure a more natural experience to the user. The system uses existing Voice-Recognition, Gesture Detection, and Text-to-Speech capabilities of handheld devices to interact with the user and guide her to a destination. Finally, our system extracts and manages semantic information about areas of the indoor environment to provide a more sophisticated navigation experience.

2 Related Work

Several kinds of sensors have been used to confront the problem of assistive indoor navigation, including: RGB cameras, depth sensors[Lee, 2011][He *et al.*, 2015], ibeacons, odometers, inertial measurement units and laser range finders. The most promising results come from sensor fusion approaches, e.g., visual inertial odometry [Lee and Dugan, 2016]. The work with CAD files in indoor localization of [Gonzalez-Aguirre *et al.*, 2014] is particularly interesting, and we have used it as a base in the retrieval system of semantic information during navigation.

3 Navigation Graph, Path Planning and Human-Computer Interaction

Our system first extracts an occupancy grid image from a CAD file or floor plan images, e.g., some buildings have a picture of the floor plan at the entrance. In the case of CAD files, our system is capable of parsing and extracting semantic information that can be used to improve the user experience while navigating. Once the occupancy information is available in the form of a grayscale image, the system creates a navigation graph using a threshold, h, that is compared to the intensity of a pixel, $I(i, j)$, at location i, j in the image. Depending on the comparison, a vertex is created. This comparison is done every g pixels. The set of vertices of the graph, extracted from the occupancy image is then,

$$V = \{[i,j]^T \mid i\%g == 0, j\%g == 0, I(i,j) > h\}, \quad (1)$$

Every vertex is encapsulated in a $node$ structure that also contains a list of all of its neighbors,

$$n_{Al} = \{m \mid m \in (N - n), \|n_v - m_v\|_2 \leq (\sqrt{2} * g)\}, \quad (2)$$

Our system uses the localization capabilities of Google's Project Tango handheld device. The device produces an Area Description File (ADF) file. We $align$ the ADF with our occupancy image, and find the corresponding transformation matrix from ADF coordinates to image coordinates. This alignment allows us to exploit the structure of the image and to codify semantic information based on that structure.

The system uses A*[Hart *et al.*, 1968] to find a path from the node in the graph that is associated to the user's current location to the node associated to the selected destination. However, since the resulting path might contain an excessive number of waypoints that will affect the user's experience during navigation, the system executes a path smoothing procedure to produce a more natural path for the user. Figure 1 illustrates the path smoothing procedure in action.

After the system is localized and the user has selected a destination, the system can proceed to run the assistive navigation module. In this stage, the system continuously checks

*This research was funded by FHWA Exploratory Advanced Research Program, award DTFH61-12-H-00002

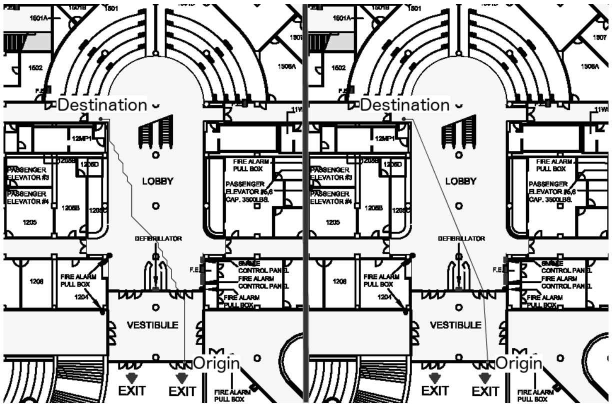

Figure 1: Path smoothing to improve user experience

	Angle to next waypoint	
Message to human user	*From*	*To*
"Go straight"	-20°	20°
"Right"	-50°	-100°
"Left"	50°	100°
"Half right"	-20°	-50°
"Half left"	20°	50°

Table 1: Guiding messages to the user during navigation

whether it can make changes in the path to accommodate the user's actions. For instance, the system checks the current distance to the destination and analyzes whether a new, better path can be generated. This analysis is based on the observation of waypoints that come later in the path.

The system also checks continuously for obstacles. It does so by using the depth information information obtained from the RGBD camera. The point cloud data is projected to 2D and this projection is compared with the contents of the occupancy image to determine whether an obstacle has been detected. Equation 3 describes how pixels in the occupancy image are marked as occupied, and how vertices are removed from the navigation graph.

$$\forall p_i \in P_I \forall p_j \in V\{\|p_i - p_j\|_2 \leq \frac{g}{2} \Rightarrow \underbrace{I(p_i.x, p_i.y) = 0}_{New\ occupied\ cells}, \quad \underbrace{V = V - p_j}_{Vertex\ removed\ from\ graph} \}, \quad (3)$$

where P_i are the 2D points (projected point cloud) and transformed to the image coordinates.

During navigation, the system provides the user with an instruction based on the angle to the next waypoint. Table 1 lists the degree values and their corresponding instructions.

4 High level semantic layer

An important functionality of the system is that it is capable of providing useful information to the user. We have implemented a spacial model based on [Gonzalez-Aguirre *et al.*, 2014] that creates a hierarchical tree of areas. The user can query for semantic information at any time. The system will perform a search on the tree and will return the label stored on the ending leaf of the search. However, the user can require more general information, which means that the system will explore back up the current branch to find additional information for the user.

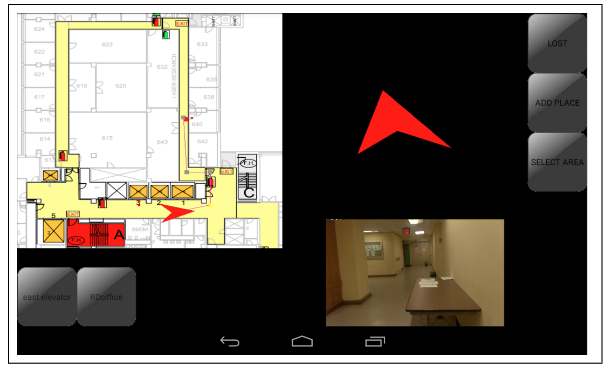

Figure 2: GUI with video for developer-only visualization

5 Conclusion

In this demo, we show a working system prototype that is capable of assisting a visually impaired user navigate indoors. We are currently improving its functionality and documenting the technical aspects of the system. Videos of our system can be found at: http://www.assistiverobot.org/ain/

Acknowledgments

We would like to thank Ivan Dryanovsky from Google, Inc. who advised us during the development stage of the system, and Mohammed Amin and Patrick Centeno for their Android programming support.

References

[Gonzalez-Aguirre *et al.*, 2014] David Gonzalez-Aguirre, Michael Vollert, Tamim Asfour, and Rudiger Dillmann. Robust real-time 6D active visual localization for humanoid robots. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 2785–2791, 2014.

[Hart *et al.*, 1968] Peter E. Hart, N.J. Nilsson, and B. Raphael. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems, Science and Cybernetics*, 4(2):100 – 107, 1968.

[He *et al.*, 2015] Hongsheng He, Yan Li, Yong Guan, and Jindong Tan. Wearable Ego-Motion Tracking for Blind Navigation in Indoor Environments. *IEEE Transactions on Automation Science and Engineering*, 12(4):1181–1190, 2015.

[Lee and Dugan, 2016] J.C. Lee and R. Dugan. Google Project Tango, 2016.

[Lee, 2011] Young Hoon Lee. RGB-D camera Based Navigation for the Visually Impaired. *RSS 2011 RGBD: Advanced Reasoning with Depth Camera Workshop*, pages 1–6, 2011.