Wearable RGBD indoor navigation system for the blind

Young Hoon Lee Gérard Medioni

Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, CA, USA
{lee126,medioni}@usc.edu

Abstract. In this paper, we present a novel wearable RGBD camera based navigation system for the visually impaired. The system is composed of a smartphone user interface, a glass-mounted RGBD camera device, a real-time navigation algorithm, and haptic feedback system. A smartphone interface provides an effective way to communicate to the system using audio and haptic feedback. In order to extract orientational information of the blind users, the navigation algorithm performs real-time 6-DOF feature based visual odometry using a glass-mounted RGBD camera as an input device. The navigation algorithm also builds a 3D voxel map of the environment and analyzes 3D traversability. A path planner of the navigation algorithm integrates information from the egomotion estimation and mapping and generates a safe and an efficient path to a waypoint delivered to the haptic feedback system. The haptic feedback system consisting of four micro-vibration motors is designed to guide the visually impaired user along the computed path and to minimize cognitive loads. The proposed system achieves real-time performance faster than 30Hz in average on a laptop, and helps the visually impaired extends the range of their activities and improve the mobility performance in a cluttered environment. The experiment results show that navigation in indoor environments with the proposed system avoids collisions successfully and improves mobility performance of the user compared to conventional and state-of-the-art mobility aid devices.

1 Introduction

Visual perception plays a very important role in everyday life, and hence visual impairment adversely affect several daily activities such as ambulating familiar and unfamiliar environments[27]. Researches have proved that vision loss severely lowers the mobility of the visually impaired [27, 4, 26]. As a result, approximately more than 30% of the blind population do not ambulate autonomously outdoors [3]. Visual impairment also increases the risk of unintentional injuries which often result in medical consequences [15, 17].

Long canes and the guide dogs have been the most popular mobility aids among the blind for navigation and collision avoidance purposes. An interview in [17] also indicated that 12% and 55% of the visually impaired interviewees

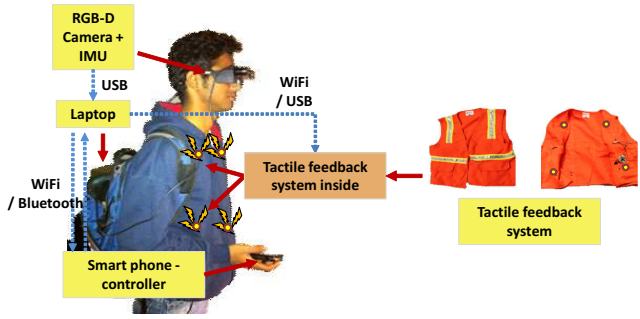

Fig. 1. Overview of the proposed system: A glass-mounted camera+IMU sensor, a smartphone user interface, and tactile interface system.

have used the guide dog and the long cane as a primary mobility aid, respectively. According to the 2011 National Health Interview Survey, 21.2 million American adults age 18 and older reported experiencing vision loss [24]. However, the guide dog has very limited availabilities as only about 1,500 individuals are estimated to graduate from a dog-guide user program [25] every year. The relatively small coverage range and a reactive nature of the long cane still accompany a high risk of collision, because the visually impaired can avoid obstacles only when they make contact with obstacles. Moreover, the long canes can cover the lower body portion of a blind user only.

To reduce risks from mobility-related activities and improve the mobility of people with vision loss, researchers have proposed Electronic Travel Aids (ETA) utilizing various types of sensors including GPS and laser sensors to replace the white cane [1, 30].

Computer vision provides rich information of surrounding environments that can be exploited for developing sophisticated applications towards real-time applications at an affordable cost, light weight, and low power consumption. Hence, the authors believe it is appropriate to transfer computer vision techniques to develop more effective wearable mobility aids for the visually impaired.

In this paper, we present a novel glass-mounted wearable RGBD camera based navigation system for the visually impaired built on the Lee's work [14]. A glass-mounted design has become a popular form factor for wearable device applications [5]. The glass-mounted platform is ideal for the wearable navigation algorithms for the following reasons. The glass-mounted design is ideal for scanning around surrounding environments as the visually impaired user traverse to reach a goal. It also extends the range of obstacle detection by detecting obstacles located at the blind sight of the long cane. The system extended a scope of navigation tasks from directing users towards the nearest open space [22, 14] to navigating to a specific destination in more robust way as presented in Section 3. In order to help the visually impaired subject navigate to a destination in an unknown environment, the proposed system analyze RGBD camera inputs to estimate 6-DOF motion of the blind subject and to perform traversability

analysis. Based on the output of the navigation system, the tactile feedback system generates a cue to guide a visually impaired user along the computed path and to alert the presence of obstacles. We present experimental results showing that the system improves the mobility performance of the blind subjects and is able to guide the blind subject to a designated goal in unknown environments without colliding to obstacles in a reasonable completion time.

The rest of this paper consists of four sections as follows. Section 2 covers recent vision based mobility devices. The proposed system with all building blocks is described in Section 3. In Section 4, experimental results are presented and we finish with conclusions and future work in Section 5.

2 Literature Review

Recently, Microsoft and Primesense introduced an affordable RGB-D camera called Kinect that provides dense 3-D information registered with an RGB image at a faster rate than that of sensors used in the past. The release of the Kinect has drawn enormous attentions among researchers and caused a surge in researches in 3D SLAM. RGBD camera based mobility aid systems are described [31, 2]. However, they were close to Proof-of-Concept platforms conducting obstacle avoidance tasks without egomotion estimation or building a global map of surroundings which is a key component of navigation system. In [18, 23], the authors proposed a stereo vision based system that estimates a camera trajectory to predict user motion and build the vicinity 3D maps. The described system, however, detects only overhead obstacles and being shoulder-mounted, does not have the advantages described in Section 1. Researchers have proposed utilizing various types of sensors to replace the white cane [1, 30]. However, such Electronic Travel Aids (ETA) designed to replace or to be attached to the long cane do not prevent head-level (chest level or higher) accidents effectively. Mobility aids based on Global Positioning System (GPS) devices [7, 16] cannot perform collision avoidance and are not specifically useful indoors.

In [22], Pradeep *et al.* proposed real-time head mounted stereo-vision based navigation system for the visually impaired. The system is known to be the first complete wearable navigation system that adopted head mounted vision system and tactile feedback prototype which brings many advantages in mobility aids applications for the blind. The head-mounted sensor design enables the visually impaired to stand and scan surroundings easier than waist- or shoulder-mounted designs that requires body rotations. The head mounted device also matches the frame of reference of the person, allowing relative position commands. The tactile interface device helps the blind users receive a resulting cue generated by the algorithms at a reasonable cognitive load through a wireless network for guidance along the safe path. The navigation system extends the coverage range and is able to detect obstacles located at the blind sight of the long cane. The experiment results indicate that the tactile vest system combined with the white cane is more efficient at alerting blind users to the presence of obstacles and helping blind subjects steer away from obstacles than the white cane alone. The

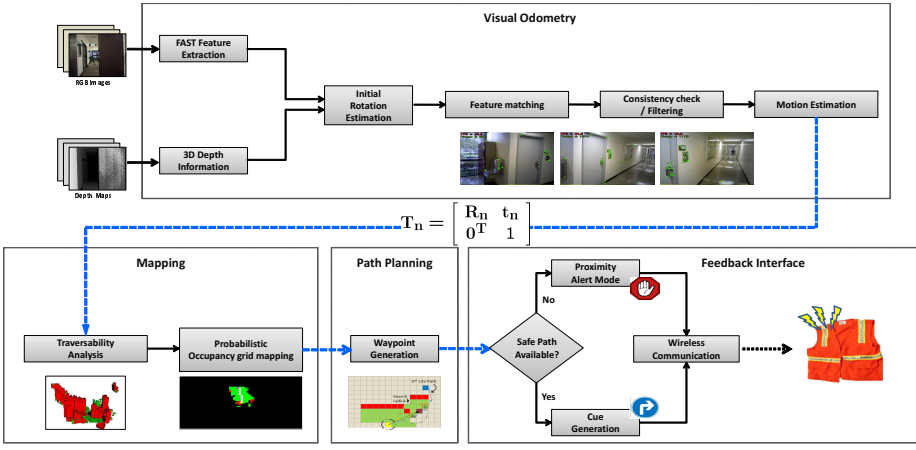

Fig. 2. Navigation algorithm overview

navigation performance was proved successful by showing trajectories generated by the proposed navigation system are the closest to the ideal trajectories from sighted subjects [22]. In [14], Lee *et. al* extended [22] and introduced a real-time wearable RGB-D camera based indoor navigation aid for the visually impaired that uses a Primesense RGB-D camera as an input device to overcome some limitations of a passive stereo camera system.

A main difference of the proposed system from [14, 22] is that a blind user is able to specify the destination for a navigation task using a smartphone application, which enables the proposed system to extend a scope of navigation tasks from a local obstacle avoidance task, that is directing a blind user towards the nearest open space without collision, to a global navigation task, guiding to a designated location. And the proposed system also adopted a probabilistic grid mapping framework which helps the system to maintain accurate and up-to-date information of surrounding environment for more robust navigation performance.

3 Navigation Algorithm

In this section, we provide an overview of the navigation algorithm as illustrated in Fig. 2. The system consists of a head-mounted RGBD camera and a vest-type interface device with four tactile feedback effectors. The wearable navigation system runs a real-time egomotion estimation, mapping, obstacle detection, and route-planning algorithm.

3.1 Initialization

The system is equipped with a small 9 DOF MEMS IMU sensor on top of RGB-D sensor for an initialization purpose. Roll and Pitch angle of the camera

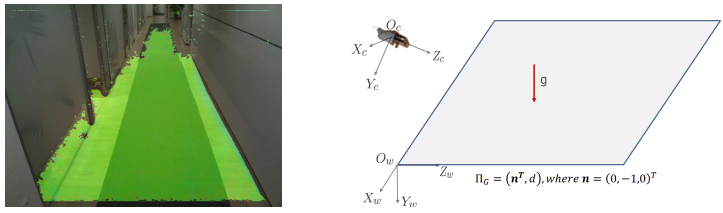

Fig. 3. (Left) Shaded areas in green color are points that have parallel to the gravity vector and used to extract the ground plane Π_G. (Right) The ground plane model is represented in the global reference frame.

orientation is initialized to be parallel to the ground plane of the world coordinate frame. Using IMU sensor readings as an orientation prior and normals of point clouds obtained using real-time normal estimation algorithm [9], we first obtain the major point clouds that have normals parallel to the gravity vector. From the set of point of clouds, we find a the plane coefficients using a RANSAC-based least square method and find the major plane $\Pi_G = (\mathbf{n}, D)$, where $\mathbf{n} = (A, B, C)^T$ is a normal vector of a plane. We then find a $\mathbf{R_0}$ such that $\mathbf{R_0} \cdot \mathbf{n} = (0, -1, 0)^T$. Yaw angle is initialized using the magnetometer reading in order to maintain an orientation with respect to the global reference frame. The pose of the blind user is represented by $\mathbf{T_n}$ where n represents the frame number.

$$\mathbf{T_0} = \begin{bmatrix} \mathbf{R_0} & \mathbf{t_0} \\ \mathbf{0^T} & 1 \end{bmatrix}, \text{ where } \mathbf{t_0} = (0, 0, 0)^T$$

The initialization step helps the visually impaired maintain consistent orientation with respect to surroundings as the blind is assumed to move on the ground plane. This is not a critical issue for offline map building applications which requires simple 3D rotations to align a model or online applications where the initial pose is always identical, for example, to a ground vehicles navigation. This initialization process, however, is essential for real-time wearable navigation system because incremental mapping methods usually build a map with respect to the camera pose of the first frame. Failures to maintain the initial orientation with respect to the world often causes the navigation system to create incorrect traversability map by converting the ground plane into obstacles. Every cell of the 2D global traversability map (GTM) is also initialized to have a uniform prior probability of $P(n) = 0.5$.

3.2 Visual Odometry

In order to maintain a sense of egocentricity of a blind user, we continuously estimate the camera pose in the world coordinate frame. We adopt the **F**ast **O**dometry from **VIS**ion (FOVIS) [12] visual odometry algorithm to achieve a real-time visual odometry. Every frame begins with an RGB and depth image capture from the camera. The RGB image is converted to grayscale and

Fig. 4. FAST corners and their correspondences from the last key-frame after refinement.

smoothed using a fixed size Gaussian kernel. Then a Gaussian pyramid is constructed to detect robust features at each Gaussian pyramid level. At each Gaussian pyramid level, FAST corners are extracted. In order to guarantee uniform distribution of features, images at each level is divided into 80×80 sub-images. the 25 strongest FAST corners are selected from each sub-image. Fast corners associated with invalid depth are discarded. Generally, the features motion in the image plane is caused by 3D rotation for continuous motion estimation applications. An algorithm proposed by Mei *et al.* [19] to compute an initial rotation by directly minimizing the sum of squared pixel errors between downsampled versions of the current and previous frames is used. This initial rotation estimation helps to constrain the search area for feature matching. 80 byte descriptor pixel patch, the last pixel omitted from 9×9, around a feature is normalized around zero mean. Then features are matched across frames using sum-of-absolute differences (SAD) as a match score. A graph of consistent feature matches is constructed using Howards approach [11]. The graph consists of vertices, a pair of feature matches, and edges that connects vertices. Rigid motions should preserve the Euclidean distance between two features matches over time in a static scene. The 3D Euclidean distance between features matches does not change drastically between consecutive frame. Hence, the set of inliers make up the maximal clique of consistent matches in a static scene. Then the maximal clique in the graph is found using a greedy algorithm [8, 11]. Initial estimation of $6-$DOF motion is calculated using Horn's absolute orientation method [10] given the inliers by minimizing the Euclidean distances between the inlier feature matches. The initial motion estimation is further refined by minimizing reprojection errors. Matches with a reprojection error above a threshold will be discarded from the inlier set and the motion estimate is refined once again to obtain the final motion estimate. Some of visual odometry results are shown in Fig. 4.

3.3 Mapping and Traversability Analysis

A 3D local voxel grid map (LVGM), a quantized representation of 3D point clouds, is generated with respect to the camera reference frame as represented in Fig. 5. Even though path planning is performed on 2D space to reduce computation complexity in a subsequent stage, traversability is analyzed in 3D space. This allows the visually impaired to prevent collisions, since the system is able

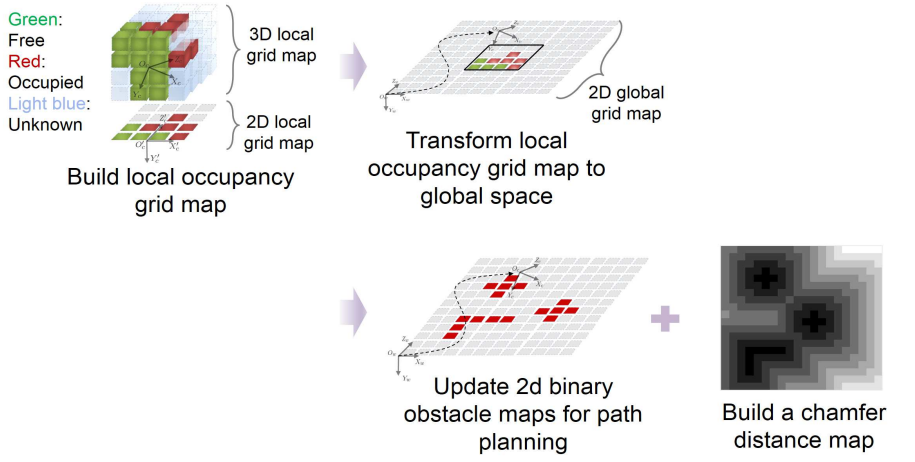

Fig. 5. Local 3D voxels is displayed with respect to the world reference frame. Local traversability is calculated O_w and O_c represent the origin of the world reference frame and the camera reference frame, respectively.

to also detect obstacles in 3D space that may be located in blind spots of the long cane.

Using the estimated camera orientation, we align the 3D LVGM in Fig. 5 with respect to the global reference frame at each frame as shown in Fig. 5. The aligned 3D LVGM is classified into occupied, free, and unknown states. Invisible voxels that are out of field of view of the camera or obstructed by another voxel are classified as unknown states. The state of the rest voxels, either occupied or free, is determined based on the number of point clouds in each voxel. Green and red voxels in a upper left corner of Fig. 5 represent free and occupied state voxels, respectively while light blue colored voxels are the voxel with unknown states. These voxels are further classified into vertical and horizontal patches which is a projection onto the local traversability map (LTM), local 2D local grid space parallel to the ground plane, centered at O'_c as shown in a upper left corner of Fig. 5. If one or more occupied voxels projected onto a cell on the LTM, the cell is registered as a vertical patch. On the other hands, cells that free voxels fall into are registered as a horizontal patch. Green cells and red cells in the bottom of a upper left corner in Fig. 5 indicate horizontal patch and vertical patches on LTM, respectively.

Glass-mounted cameras for the navigation often faces an spatial configuration where the most of point clouds from a RGBD sensor are located beyond a certain distance as the goal of the system is to guide the visually impaired to steer the subject away from possible collision threats such as walls and obstacles which act as a reliable source when estimating motion of the camera. Hence, depth values from the glass-mounted platform tend to be quite noisy since the noise characteristic of the RGB-D sensors is proportional to the quadratic of distance to a point [21] We have empirically found out that a single vertical patch with-

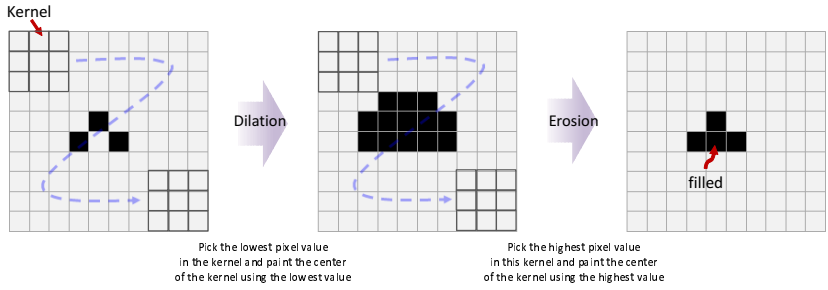

Fig. 6. A filtering process using erosion and dilation. Black and white cells represent vertical and horizontal patches, respectively.

out neighboring vertical patches or a isolated horizontal patch surrounded by neighboring vertical patches generally result from noisy sensor reading. Hence, we apply a simple filtering algorithm using 2D image processing techniques, erosion and dilation, for noise reductions on LTM. Fig. 6 shows a simple example how this erosion and dilation process fills a hole on a local grid map that is likely to be misinterpretation of the environment introduced by sensor noise.

Then the vertical and horizontal patches in the LTM are transformed to the world reference frame to update the occupancy probabilities of the 2D GTM as displayed in 5. The occupancy probability of each cell of GTM is updated by occupancy grid mapping rule suggested by Moravec[20].

$$P(n|z_{1:t}) = \left[1 + \frac{1 - P(n|z_t)}{P(n|z_t)} \frac{1 - P(n|z_{1:t-1})}{P(n|z_{1:t-1})} \frac{P(n)}{1 - P(n)}\right]^{-1} \tag{1}$$

For efficient computation in updating, we use $logOdds()$ notation which can be directly converted into probabilistic values when needed. As stated in [28], this notation replaces multiplication with addition which is more efficient in computation.

$$L(n|z_{1:t}) = L(n|z_{1:t-1}) + L(n|z_t) \tag{2}$$

One disadvantage of the update policy represented in (2) is that it requires as many observation as had been integrated before to obtain the current state. In order for the system to respond to dynamic changes in the environment immediately and overcome the overconfidence in the map, the upper and lower bound, l_{max} and l_{min}, of the *logodds* values is enforced using clamping update policy proposed by Yguel [29].

$$L(n|z_{1:t}) = max(min(L(n|z_{1:t-1}) + L(n|z_t), l_{max}), l_{min}) \tag{3}$$

Cells on the probabilistic global traversability map whose occupancy probability is lower than $P(l_{min})$ are registered as traversable while cells with higher occupancy probability than $P(l_{max})$ are registered as non-traversable, where $P(l_{min})$ and $P(l_{max})$ are corresponding probability value of l_{min} and l_{max}, respectively. As stated 3.1, all cells are initialized to have unknown state. Once the state of a

cell changed to either traversable or non-traversable from unknown, cells whose occupancy probability value falls between $P(l_{max})$ and $P(l_{min})$, are registered as unknown areas.

3.4 Path planning

The blind subject is assumed to travel on a ground plane. Therefore, the path planning is performed on the global 2D traversability map for computational efficiency. We also build and update a chamfer distance array that stores distance to the closest obstacle corresponding to the global 2D traversability map which is used to verify a direct measure of risk of collision instantaneously. The shortest path is produced using the D* Lite Algorithm as suggested by [13] from the current location to the destination. D* algorithm can handle dynamic changes of the surrounding very efficiently. However, subtle changes in the map sometimes result in changes of the generated path in high frequency, which confuses the blind subject. The shortest path produced is a set of cells, DL, on the grid map which connects the current location to the destination without using untraversable cells. Instead of directly following DL, we generate a reliable waypoint, W, that are confirmed to be traversable and located at least some distance from obstacles. In order to generate a waypoint, we find the furthest point in DL which is directly visible and traversable from the current location, D_1, and another cell, D_2, located furthest in the set DL which is directly visible and traversable from D_1. Then search the neighborhood of D_1 limited by a predefined radius and find a cell that minimize the cost function, f, given in (5). $C_{i,j}$ and $CF_{i,j}$ represent a cell in i^{th} row and j^{th} column and Chamfer distance value of $C_{i,j}$, respectively. $d(C_1, C_2)$ indicates distance between two cell, C_1 and C_2.

$$W = \underset{i,j}{\operatorname{argmin}} f_{i,j} \ , \tag{4}$$

$$\text{where } f_{i,j} = w_1 \times [d(C_{i,j}, D_1) + d(C_{i,j}, D_2)] + w_2 \times CF_{i,j}$$

The temporary waypoint is updated when there is motion or rotation greater than a certain threshold , when the current waypoint is not visible from the

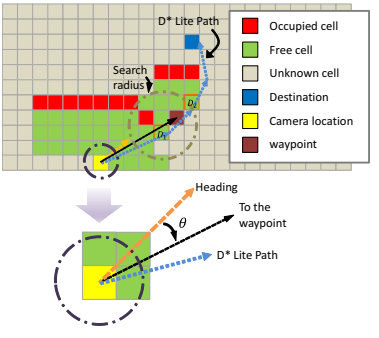

Fig. 7. An example of a waypoint update scenario.

current location, or the blind subject arrives the vicinity of the waypoint to avoid frequent updates of the generated path and redundant computation.

3.5 Cue generation and Haptic feedback

The tactile feedback instead of audio feedback was used in order to reduce the amount of cognitive loads from hearing, which most blind users rely on for various other tasks. This system has intuitive 4 different cues for a user as followings:

- Straight (no tactile sensors on)
- Stop & Scan: Proximity Alert (all tactile sensors on)
- Turn left (top-left sensor on)
- Turn right (top-right sensor on)

The cue is generated simply based on a relative angle between a direction obtained by drawing a line from the current position to the waypoint and the current heading. In order to prevent frequent changes in cues which often causes confusions from the visually impaired subject, we utilize a hysteresis loop equation that takes the current status of a cue into account as well as a rotation angle required to reach the waypoint, W, as follows.

$$\text{New Cue} = \begin{cases} \text{Turn Left if} & -180^\circ < \theta \leq -\theta_2, \\ \text{Curr. Cue if} & -\theta_2 < \theta < -\theta_1, \\ \text{Straight if} & -\theta_1 \leq \theta \leq \theta_1, \\ \text{Curr. Cue if} & \theta_1 < \theta < \theta_2, \\ \text{Turn Right if} & \theta_2 \leq \theta < 180^\circ. \end{cases}$$

For example, let us assume the rotation angle, θ, falls somewhere between θ_1 and θ_2. If the current cue is **Straight**, a new cue will still be **Straight**. On the other hands, if the current cue is **Turn right**, a new cue will be **Turn right** until a subject completes rotation so that $\theta < \theta_1$. In our application, θ_1 and θ_2 are defined as 5° and 15°, respectively. The generated cue is transmitted through the wireless transmitter to the vest interface that has four vibration motors as provided in Fig. 1. Each vibration motor delivers a specific cue such as turn left, turn right, and so on.

4 Experiments

We present the experimental details and results of mobility experiments of the proposed system . Our navigation system runs at 31 Hz on average on the following configuration.

- CPU: Intel(R) Core(TM) i7-3630QM CPU @ 2.40GHz
- RAM: 8.00 GB
- OS: Windows 7 - 64 bit

The breakdown of processing timing is represented in Fig. 8.

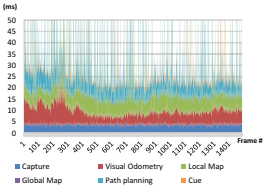

Fig. 8. Timings of each process over 1500 frames.

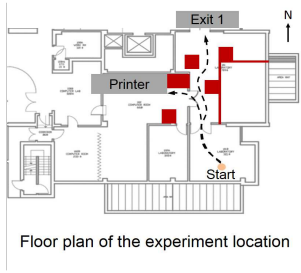

Fig. 9. A floor plan of the mobility experiments. Red boxes and lines indicates obstacles that are not reflected in the floor plan.

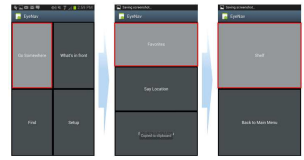

Fig. 10. Snapshots of smartphone interface for a blind user

4.1 Mobility experiments (vs. conventional aids)

These experiments are designed to test an ability of the system to guide a blind user in a cluttered environment. In this environment, there exist many obstacles such as tables, chairs, cublcle walls, and a fridge that are not present on the floorplan, which is a primary reason that the system does not rely entirely on the floorplan for mapping and planning and performs a real-time probabilistic map update. 4 subjects are blind-folded and asked to reach a destination in 3 different scenarios, 3 times per each scenario. In the first scenario, the subjects are asked to travel to a destination whose location is known using the white cane and provided the initial orientation towards the destination. This scenario is very similar to what one experiences in an office or at home where you spend most of your time. In the second scenario, the subjects are asked to travel to a destination whose location is known using the white cane and does not know the initial orientation towards the destination. This scenario is designed to simulate a situation often happening in a daily basis even in a very familiar environment when a subject loses a sense of the orientation or gets confused by some changes in the spatial configuration. The last scenario has the same configuration to the second scenario except that the subjects are provided with the proposed navigation system in addition to the white cane. For the last scenario, a database that contains location of interests in the building is assumed to be available at the beginning for all experiments. The database only stores 2D location of interests of a building which is very concise and can be easily appended on top of existing maps, an indoor google map for example. We extract the location of rooms and some points of interests such as elevators, exits, printers, and a fridge from using floor plan as shown in Fig. 9 aligning them to a grid where the north points up and store them in a database. You provide the system with the destination and the current location, for example to a printer room from Room 101C, to start a navigation task.

Fig. 11. The average elapsed time to complete a navigation task for different scenarios.

It is obvious that you will get to a destination the fastest when you know the surrounding environments and orientation very well which is not always the case when you travel to different places. One thing to note is that losing a sense of orientation can decrease the mobility performance by about a half in terms of the completion time to reach a goal even in familiar environments which will get probably worse in unfamiliar environments. As you can see in Fig 11, the proposed system improved the mobility performance by 57.77% on average, 79.0% and 36.5% for each destination. This results from an ability of the system that can initialize the orientation using magnetometer reading and help the subjects to align themselves to the existing map.

4.2 Mobility experiments (vs. the state-of-the-art ETA)

It is difficult to directly compare performance of the navigation systems to other existing mobility aids than the white canes quantitatively since there are no standard courses to complete nor metrics to measure. However, walking speed is a good indicator how efficient the mobility aids is in helping the blind subject navigate around. A recent research reported $0.2m/s$ of walking speed of blind-folded subject in a non-cluttered environment [6] which can serve as a baseline for a performance comparison purpose. Hence, we set up an testing environment to perform a quantitative comparison, which contains 3 sharp curves and the width of the narrowest corridor was 0.9m. 3 blind-folded subjects were asked to follow cues from the navigation system to reach a goal, an exit door of a building whose location is unknown in advance to the subject, in an unknown environment. This test is to simulate to a case where a blind user arrives a building (unknown environments) and has a limited information such as the approximate current location, an entrance for example, and the destination in the building. Total travel distance was approximately $25m$ and the course includes obstacles including doors, chairs, as shown in Fig. 12. As shown in Fig. 12, the blind-folded subjects were able to finish the course successfully without colliding to obstacles.

The subjects completed the course in 89 seconds on average resulting in $0.28m/s$ of walking speed in a non-cluttered environment, which proves the navigation system works very effectively in an unknown environment. Note is that the ability of the proposed system to guide blind subjects in an unknown environment indicates that the system is scalable to multiple buildings as a building can be easily identified by either building name or GPS information from your smartphone. Since the system builds the map of surrounding environments online, the system does not need to know all structures such as walls and so on in advance and one can generate a location database of a building and share it easily on the web.

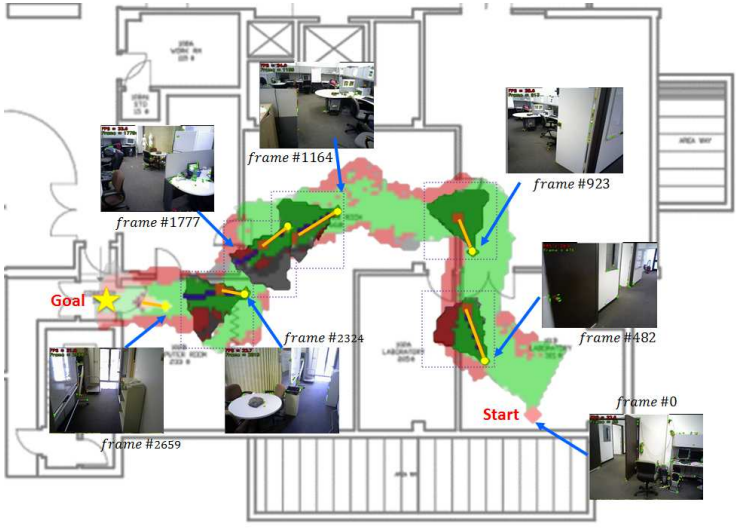

Fig. 12. A 2D GTM map generate when traveling in a non-cluttered environment: Green and red regions are horizontal and vertical patches, respectively. Yellow dots in darker areas represent a pose of the blind subject at n^{th} frame. Orange bars in darker indicate direction to the temporary goals.

5 Conclusion and Future works

We have presented an integrated framework using RGBD sensor as an input to operate real-time navigation algorithms for an application in a mobility aids for the blind. We have also shown that an intelligent set of tactile feedback cues can guide visually impaired subjects through obstacles at a reasonable cognitive load. The mobility experiment results suggest that the system improves the mobility performance of the blind-folded users by about 58% than the blind-folded users did using their white canes when the orientation is unknown in the beginning of the navigation tasks. We also compared performance of the system

in terms of the completion time of a navigation task with respect to one of the state of state-of-the-art technique [6] and we proved that the effectiveness of the system. We are planning on adopting a loop closure algorithm because in many cases people visit a few places multiple times and a loop closure detection helps to mitigate error accumulation problem and to optimize pose estimation when you revisit a place. We have encountered severe image blurs or poor feature matching led by rapid head rotations sometimes, which caused inaccurate visual odometry results. Inertia-visual sensor fusion has been proven to provide much more robust tracking results due to complementary characteristics of an IMU sensor and cameras. For this reason, we plan to extend the range of an IMU sensor integration that is utilized only at the initialization stage under the proposed system. Finally, experiments with the real visually impaired subjects in various environments are planned in order to evaluate the effectiveness of the navigation system.

Acknowledgement

This research was made possible by a cooperative agreement that was awarded and administered by the U.S. Army Medical Research & Materiel Command (USAMRMC) and the Telemedicine & Advanced Technology Research Center (TATRC), at Fort Detrick, MD under Contract Number: W81XWH-10-2-0076. The authors would like to acknowledge Nii Mante and Prof. Weiland in Biomedical Engineering department at USC for contributions to development of the smartphone application.

References

1. Borenstein, J., Ulrich, I.: The GuideCane - A computerized travel aid for the active guidance of blind pedestrians. In: IEEE Int. Conf. on Robotics and Automation. pp. 1283–1288 (1997)
2. Brock, M., Kristensson, P.: Supporting blind navigation using depth sensing and sonification. In: Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication. pp. 255–258 (2013)
3. Clark-Carter, D., Heyes, A., Howarth, C.: The efficiency and walking speed of visually impaired people. Ergonomics 29(6), 779–789 (1986)
4. Golledge, R.G., Marston, J.R., Costanzo, C.M.: Attituds of visually imparied persons towards the use of public transportation. Journal of Visual Impairment Blindness 91(5), 446–459 (1997)
5. Google: Google glass (2013), http://www.google.com/glass
6. Guerrero, L., Vasquez, F., Ochoa, S.: An indoor navigation system for the visually impaired. Sensors 12(6), 8236–8258 (2012)
7. Helal, A., Moore, S., Ramachandran, B.: Drishti: An integrated navigation system for visually impaired and disabled. In: Wearable Computers, 2001. Proceedings. Fifth International Symposium on. pp. 149–156. IEEE (2001)
8. Hirschmuller, H., Innocent, P., Garibaldi, J.: Fast, unconstrained camera motion estimation from stereo without tracking and robust statistics. In: Control, Automation, Robotics and Vision, 2002. ICARCV 2002. 7th International Conference on. vol. 2, pp. 1099–1104. IEEE (2002)

9. Holzer, S., Rusu, R., Dixon, M., Gedikli, S., Navab, N.: Adaptive neighborhood selection for real-time surface normal estimation from organized point cloud data using integral images. In: Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on. pp. 2684–2689. IEEE (2012)
10. Horn, B.: Closed-form solution of absolute orientation using unit quaternions. JOSA A 4(4), 629–642 (1987)
11. Howard, A.: Real-time stereo visual odometry for autonomous ground vehicles. In: Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on. pp. 3946–3952. IEEE (2008)
12. Huang, A., Bachrach, A., Henry, P., Krainin, M., Maturana, D., Fox, D., Roy, N.: Visual odometry and mapping for autonomous flight using an rgb-d camera. In: International Symposium on Robotics Research (ISRR). pp. 1–16 (2011)
13. Koenig, S., Likhachev, M.: Fast replanning for navigation in unknown terrain. IEEE Transactions on Robotics 3(21), 354–363 (2005)
14. Lee, Y., Medioni, G.: A RGB-D camera based navigation for the visually impaired. In: RSS 2011 RGB-D: Advanced Reasoning with Depth Camera Workshop. pp. 1–6 (2011)
15. Legood, R., Scuffham, P., Cryer, C.: Are we blind to injuries in the visually impaired? a review of the literature. Injury Prevention 8(2), 155–160 (2002)
16. Loomis, J., Golledge, R., Klatzky, R.: Gps-based navigation systems for the visually impaired. Fundamentals of Wearable Computers and Augmented Reality 429(46) (2001)
17. Manduchi, R., Kurniawan, S.: Mobility-related accidents experienced by people with visual impairment. Insight: Research & Practice in Visual Impairment & Blindness 4(2), 44–54 (2011)
18. Martinez, J., Ruiz, F., et al.: Stereo-based aerial obstacle detection for the visually impaired. In: Workshop on Computer Vision Applications for the Visually Impaired (2008)
19. Mei, C., Sibley, G., Cummins, M., Newman, P., Reid, I.: A constant-time efficient stereo slam system. In: BMVC. pp. 1–11 (2009)
20. Moravec, H., Elfes, A.: High resolution maps from wide angle sonar. In: Robotics and Automation. Proceedings. 1985 IEEE International Conference on. vol. 2, pp. 116–121. IEEE (1985)
21. Nguyen, C., Izadi, S., Lovell, D.: Modeling kinect sensor noise for improved 3d reconstruction and tracking. In: 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on. pp. 524–530. IEEE (2012)
22. Pradeep, V., Medioni, G., Weiland, J.: Robot vision for the visually impaired. In: Computer Vision Applications for the Visually Impaired. pp. 15–22 (2010)
23. Sáez, J., Escolano, F.: 6dof entropy minimization slam for stereo-based wearable devices. Computer Vision and Image Understanding 115(2), 270–285 (2011)
24. Schiller, J., Peregoy, J.: Provisional report: Summary health statistics for u.s. adults: National health interview survey, 2011. National Center for Health Statistics 10(256) (2012)
25. news service, J.: Demographics update: Alternate estimate of the number of guide dog users. Journal of Visual Impairment Blindness 89(2), 4–6 (1995)
26. Turano, K., Broman, A., Bandeen-Roche, K., Munoz, B., Rubin, G., West, S., TEAM, S.P.: Association of visual field loss and mobility performance in older adults: Salisbury eye evaluation study. Optometry & Vision Science 81(5), 298–307 (2004)

27. West, S., Rubin, G., Broman, A., Munoz, B., Bandeen-Roche, K., Turano, K.: How does visual impairment affect performance on tasks of everyday life?: The see project. Archives of Ophthalmology 120(6), 774–780 (2002)
28. Wurm, K., Hornung, A., Bennewitz, M., Stachniss, C., Burgard, W.: Octomap: A probabilistic, flexible, and compact 3d map representation for robotic systems. In: Proc. of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation. vol. 2 (2010)
29. Yguel, M., Aycard, O., Laugier, C.: Update policy of dense maps: Efficient algorithms and sparse representation. In: Field and Service Robotics. pp. 23–33. Springer (2008)
30. Yuan, D., Manduchi, R.: Dynamic environment exploration using a virtual white cane. In: IEEE Conf. on Computer Vision and Pattern Recognition. pp. 243–249 (2005)
31. Zllner, M., Huber, S., Jetter, H., Reiterer, H.: Navi a proof-of-concept of a mobile navigational aid for visually impaired based on the microsoft kinect. In: Human-Computer Interaction INTERACT 2011. pp. 584–587 (2011)