Assistive Robotics for Empowering Humans with Visual Impairments to Independently Perform Day-to-day Tasks

Doctoral Consortium

Shivendra Agrawal
University of Colorado Boulder
shivendra.agrawal@colorado.edu

ABSTRACT

The ability to perform common day-to-day tasks is essential for an independent lifestyle. However, many crucial tasks are unaddressed for blind or visually impaired (BVI) people with the current solutions. Our research goal aims to provide technical solutions to such problems to help support more autonomy for BVI people. Through this work, we present a proof-of-concept socially assistive robotic cane that can assist with 1) a navigation task which is finding a socially preferred seat in unknown public places and guiding the users toward it, 2) a manipulation task which is locating and retrieving the desired product from a grocery store shelf. We evaluated our system in an initial pilot study with sighted blindfolded testers, with encouraging results that show the system's potential to provide purposeful and effective navigation guidance optimizing for users' convenience, privacy, and intimacy while increasing their confidence in independent navigation. Another study we ran showed the system's success in locating and providing effective fine-grain manipulation guidance to retrieve desired products with novice users while eliciting a positive user experience.

KEYWORDS

Assistive Robotics; Computer Vision; Planner; Manipulation Guidance; Human-Robot Interaction; Markov Decision Process

ACM Reference Format:
Shivendra Agrawal. 2023. Assistive Robotics for Empowering Humans with Visual Impairments to Independently Perform Day-to-day Tasks: Doctoral Consortium. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

Figure 1: The system is comprised of a robotic cane equipped with RealSense D455 and T265 cameras. The system is powered through a laptop in a backpack. *Left:* Utilizing haptic feedback for navigation guidance. *Right:* Utilizing audio cues for manipulation guidance.

1 MOTIVATION

Service dogs and white canes are the most common aids used by BVI individuals. Service dogs can cost upwards of $50,000 to train, and incur $1,200 on average in annual care costs. White canes do well in tracing along walls, curbs, and entrances, but they offer very little utility in social contexts (including finding seating in public places and avoiding contact with others in crowded environments). Wang et al. [20] reported that finding available seating in public areas as one of the most important mobility tasks not addressed by current solutions, rated as '5 out of 5' in terms of importance on a survey of BVI individuals. Staats and Groot [18] showed that people choose to seat themselves with respect to others in a way that optimizes *intimacy* and *privacy*. With current technology, BVI people are not able to independently locate seating, let alone take into account the nuances of seat choice available to sighted people.

BVI shoppers have also indicated they are not willing to use store staffers for shopping for items that require discretion such as medicine and personal hygiene items [1, 10, 11]. Our work seeks to alleviate the dependence on the sighted guide's availability and to attenuate the loss of privacy. We ground our manipulation work within the grocery store domain both for immediate broader impact and because it contains dense concentrations of similar items shelved together that make for very poor tactile differentiability.

Here we summarize our work developing an autonomous robotic cane that enables goal-based navigation in unknown, indoor environments while optimizing for convenience, privacy, and intimacy as our prior work to support an important mobility task, and also describe our second work in which the same robotic system provides fine-grain manipulation guidance with a novel visual product locator algorithm designed for use in grocery stores and a novel planner that autonomously issues verbal manipulation guidance commands to guide the user during product retrieval.

2 CURRENT WORK

2.1 Social navigational guidance

Design: We designed a robotic cane (Fig 1) composed of RealSense cameras and vibrational motors mounted on a cane, with a backpack-worn laptop to run perception and navigation algorithms that can locate socially preferred seats (i.e. the seats that have higher privacy and lower intimacy) and guide the user to it. The software

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

system follows a serial architecture with perception, planning, and conveyance. The system maps the environment and identifies people and things in the environment using onboard SLAM, object detection, and data association. Planning involves two steps: Goal selection and Path planning. Each goal (chairs in the case of social seating selection) is scored by an objective function that accounts for levels of *convenience*, *privacy*, and *intimacy*.

Social goal scoring: We developed a novel anchor scoring algorithm for finding relative anchor scores for each chair. The key insight was to over-measure contiguous obstacles which are similar enough to walls and can be used to quantify the notion of privacy. We also developed a linear cost model to assign a proximity score to all the seats accounting for intimacy as closeness to other humans, convenience in terms of goal distance, and closeness to items such as backpacks or laptops, as these can indicate an unavailable goal. The system then used RRT* to find a path to the goal. This motion plan is conveyed to the user through auditory and haptic modalities. Auditory feedback is used to provide a plan overview, vibrotactile feedback (via two coin motors such that the user's thumb is placed between them [14]) is used to provide online navigational guidance.

Pilot Study: The pilot was successful in validating the system with 6 novice users each performing 6 trials. Users of the system had a 100% success rate at finding any seat and an 83.3% success rate at finding the socially-preferred seat. They rated the system highly on confidence in navigation and goal-finding, as well as verbal overview helpfulness. We presented this work [3] at IROS 2022.

2.2 Fine-grain manipulation guidance

Design: Grocery shopping primarily consists of three main subtasks: navigation, product retrieval, and product examination. Our current work called ShelfHelp [2] focuses on product retrieval. ShelfHelp extends the capability of our robotic cane [3] discussed earlier. The area of assistive navigation has been extensively researched [3, 8, 9, 15–17, 19, 21] and it is only prudent to utilize the sensing and compute of these existing systems.

Product location: ShelfHelp employs a novel 2-stage computer vision pipeline to search and locate desired grocery products assuming that the user is in front of the correct shelf. It requires the user to have a single image of the product that they want to find. In the first stage, we use a YoloV5 network that gives us the most likely bounding boxes to contain *any* product. In the second stage, we take these regions and compare each region against the image of the desired product by extracting features from both. ShelfHelp locates the product in real-time and doesn't require retraining or any environmental augmentation unlike some prior work [5, 7].

Fine-grain manipulation guidance: We also developed a novel fine-grain manipulation guidance system that optimizes for guide time and the # of commands without compromising legibility. We formulate the guidance problem as an MDP (S, A, T, R) (Fig 2) where a state corresponds to the remaining distance to the goal. A is the set of discrete verbal commands such as *"Move 6 inches to the left"*. We collected a dataset mapping verbal commands from a command-set, recording participants' net hand movements upon reacting to that command (Fig 2). We fitted Gaussians to characterize the net movement which would give us the transition probability matrix T. The reward function R encourages reaching the target and discourages issuing superfluous commands. It also discourages a sequence

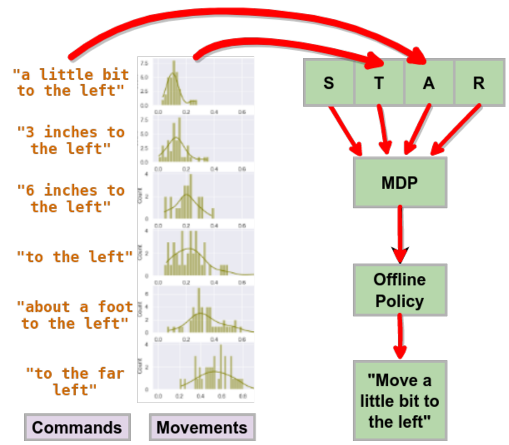

Figure 2: *(Left to right)* A sample of discrete commands. The movement (in meters) each command caused. A policy is learned offline that can be used across reaching tasks.

of commands that could be illegible by penalizing axis changes.

Result: We did a system-level validation with 15 blindfolded participants. At this stage of design and technical readiness, we follow prior work in performing preliminary validations using blindfolded users [4, 6, 12, 13, 17, 19] as a precursor to engaging with the BVI community. We compared our guidance algorithm with a baseline algorithm and a human over a video call. The product detection failed **21/150** times in locating the desired product mainly due to the product not being in frame. This led the system to find the most visually similar product it can find. Our planner guided the users to the correct product **75/75** times. The users still picked up the wrong item on 6/75 times as they picked the adjacent item as the items were closely placed. Our planner performed statistically better in terms of # of commands and guide time compared to the baseline and was on par with the human caller. The study also showed positive feedback for qualitative metrics such as ease of use, confidence, mental demand, temporal demand, frustration, intelligence, and competence. ShelfHelp [2] is going to be presented at AAMAS 23.

3 FUTURE WORK

The issues presented by an unstructured world require research in sensing, planning, and conveyance which are all core robotics problems at heart. We are now starting to work on building algorithms that can provide task-oriented navigational guidance to users, including directing them to desired products by locating the correct shelf in a store without needing environmental augmentation. We aim to explore 1) The capability to map unknown stores and semantically label and segment the map into task-centric regions such as shelves, aisles, produce-section, exits, checkout areas, etc. based on the spatial distribution of grocery items and landmarks, 2) Using behavior trees to inform situationally appropriate guidance for tasks (e.g., grocery shopping), 3) Motion planning techniques that can optimize for metrics such as user convenience, safety, total time, # of turns, and 4) Multi-modal conveyance methods that communicate navigational and richer semantic information efficiently through an interplay of audio and vibrotactile mediums.

REFERENCES

[1] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy concerns and behaviors of people with visual impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3523–3532.

[2] Anonymized. 2023. Anonymized. In *22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

[3] Anonymized. In press. Anonymized. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. -

[4] Aline Darc dos Santos, Fausto Orsi Medola, Milton José Cinelli, Alejandro Rafael Garcia Ramirez, and Frode Eika Sandnes. 2020. Are electronic white canes better than traditional canes? A comparative study with blind and blindfolded participants. *Universal Access in the Information Society* 20, 1 (2020). https://doi.org/10.1007/s10209-020-00712-z

[5] Chia-Hui Feng, Ju-Yen Hsieh, Yu-Hsiu Hung, Chung-Jen Chen, and Cheng-Hung Chen. 2020. Research on the Visually Impaired Individuals Shopping with Artificial Intelligence Image Recognition Assistance. In *International Conference on Human-Computer Interaction*. Springer, 518–531.

[6] A Jin Fukasawa and Kazusihge Magatani. [n.d.]. A navigation system for the visually impaired an intelligent white cane. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*

[7] Chaitanya P Gharpure and Vladimir A Kulyukin. 2008. Robot-assisted shopping for the blind: issues in spatial cognition and product selection. *Intelligent Service Robotics* 1, 3 (2008), 237–251.

[8] João Guerreiro, Daisuke Sato, Saki Asakawa, Huixu Dong, Kris M. Kitani, and Chieko Asakawa. 2019. CaBot: Designing and Evaluating an Autonomous Navigation Robot for Blind People. In *The 21st ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA). NY, USA. https://doi.org/10.1145/3308561.3353771

[9] Robert K. Katzschmann, Brandon Araki, and Daniela Rus. 2018. Safe Local Navigation for Visually Impaired Users With a Time-of-Flight and Haptic Feedback Device. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 3 (2018), 583–593. https://doi.org/10.1109/TNSRE.2018.2800665

[10] Vladimir Kulyukin, Chaitanya Gharpure, and John Nicholson. 2005. Robocart: Toward robot-assisted navigation of grocery stores by the visually impaired. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2845–2850.

[11] Vladimir A Kulyukin and Chaitanya Gharpure. 2006. Ergonomics-for-one in a robotic shopping cart for the blind. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 142–149.

[12] Ryuta Okazaki and Hiroyuki Kajimoto. 2014. Perceived distance from hitting with a stick is altered by overlapping vibration to holding hand. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 1903–1908.

[13] Emily E O'Brien, Aaron A Mohtar, Laura E Diment, and Karen J Reynolds. 2014. A detachable electronic device for use with a long white cane to assist with mobility. *Assistive Technology* 26, 4 (2014), 219–226.

[14] C. A. Perez, C. A. Holzmann, and H. E. Jaeschke. [n.d.]. Two-point vibrotactile discrimination related to parameters of pulse burst stimulus.

[15] Santiago Real and Alvaro Araujo. 2019. Navigation Systems for the Blind and Visually Impaired: Past Work, Challenges, and Open Problems. *Sensors* (2019). https://doi.org/10.3390/s19153404

[16] Manaswi Saha, Alexander J Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2019. Closing the gap: Designing for the last-few-meters wayfinding problem for people with visual impairments. In *The 21st international acm sigaccess conference on computers and accessibility*. 222–235.

[17] Patrick Slade, Arjun Tambe, and Mykel J Kochenderfer. 2021. Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision. *Science Robotics* 6, 59 (2021).

[18] Henk Staats and Piet Groot. 2019. Seat Choice in a Crowded Café: Effects of Eye Contact, Distance, and Anchoring. *Frontiers in Psychology* 10 (2019), 331. https://doi.org/10.3389/fpsyg.2019.00331

[19] Andreas Wachaja, Pratik Agarwal, Mathias Zink, Miguel Reyes Adame, Knut Möller, and Wolfram Burgard. 2017. Navigating blind people with walking impairments using a smart walker. *Autonomous Robots* 41, 3 (2017).

[20] Hsueh-Cheng Wang, Robert K. Katzschmann, Santani Teng, Brandon Araki, Laura Giarré, and Daniela Rus. 2017. Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. https://doi.org/10.1109/ICRA.2017.7989772

[21] Anxing Xiao, Wenzhe Tong, Lizhi Yang, Jun Zeng, Zhongyu Li, and Koushil Sreenath. 2021. Robotic Guide Dog: Leading a Human with Leash-Guided Hybrid Physical Interaction. arXiv:2103.14300 [cs.RO]