Clew3D: Automated Generation of O&M Instructions Using LIDAR-Equipped Smartphones

Hwei-Shin Harriman
hwei.shin.harriman@gmail.com
Franklin W. Olin College of Engineering
Needham, Massachusetts, USA

Dragan Ahmetovic
Università degli Studi di Milano
Milan, Milan, Italy
dragan.ahmetovic@unimi.it

Sergio Mascetti
Università degli Studi di Milano
Milan, Milan, Italy
sergio.mascetti@unimi.it

Darren Moyle
Vision Australia
Melbourne, Australia
Darren.Moyle@visionaustralia.org

Michael Evans
Vision Australia
Melbourne, Australia
Michael.Evans@visionaustralia.org

Paul Ruvolo
Franklin W. Olin College of Engineering
Needham, USA
paul.ruvolo@olin.edu

ABSTRACT

Certified orientation and mobility specialists (COMS) work with clients who are blind or visually impaired (BVI) to help them travel independently with confidence. Part of this process involves creating a narrative description of a route and using specific techniques to help the client internalize it. We focus on the problem of automatically generating a narrative description of an indoor route based on a recording from a smartphone. These automatically generated narrations could be used in cases where a COMS is not available or to enable clients to independently practice routes that were originally learned with the help of a COMS. Specifically, we introduce Clew3D, a mobile app that leverages LIDAR-equipped iOS devices to identify orientation and mobility (O&M) landmarks and their relative location along a recorded route. The identified landmarks are then used to provide a spoken narration modeled after traditional O&M techniques. Our solution is co-designed with COMS and uses methods and language that they employ when creating route narrations for their clients. In addition to presenting Clew3D, we report the results of an analysis conducted with COMS regarding techniques and terminology used in traditional, in-person O&M instruction. We also discuss challenges posed by vision-based systems to achieve automatic narrations that are reliable. Finally, we provide an example of an automatically generated route description and compare it with the same route provided by a COMS.

CCS CONCEPTS

• **Human-centered computing → Accessibility systems and tools**.

KEYWORDS

computer vision, guided navigation, object classification, orientation and mobility

ASSETS '21, October 18–22, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8306-6/21/10.
https://doi.org/10.1145/3441852.3476564

ACM Reference Format:
Hwei-Shin Harriman, Dragan Ahmetovic, Sergio Mascetti, Darren Moyle, Michael Evans, and Paul Ruvolo. 2021. Clew3D: Automated Generation of O&M Instructions Using LIDAR-Equipped Smartphones. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21), October 18–22, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3441852.3476564

1 INTRODUCTION

For individuals who are blind or visually impaired (BVI), the ability to learn navigation routes is critical for their day-to-day activities. Depending on the individual's comfort and skill level navigating their environment [2], even tiny mistakes can lead to veering off course and missing key landmarks along the route [3]. Because of this, certified orientation and mobility specialists (COMS) work with individuals who are BVI to create routes that can be internalized by the individual using specific techniques and language [11].

Our existing system, Clew [12], is an iPhone app available on the App Store that allows COMS or individuals who are BVI to record and save routes, which can then be loaded and navigated using automatically generated haptic, speech, and auditory guidance. The current implementation has only basic functionality in route narration, providing instructions such as "slight left" or "slight right" as well as distances to intermediate waypoints. The current approach is unable to provide contextual information about the environment, which is important for successful navigation [11]. The addition of LIDAR (Light Detection and Ranging) to iOS devices released in the year 2020 introduces the potential for more robust implementations. Specifically, the combination of a LIDAR sensor with computer vision algorithms provides more insight about the immediate surroundings, including accurate and immediate surface detection, depth sensing, and environmental mesh creation.

We are collaborating with one professor of O&M and two professional O&M instructors to determine what contextual information about a given route is most useful to individuals who are BVI. Though each COMS has their own particular style and routes are often customized to individual students, there are several core techniques and terms that COMS are trained to use to make routes easier for their students to learn and follow. These methods help the person who is BVI to internalize the route by informing them of various contextual features of their environment and providing descriptions that conjure mental images of the route. For instance, a COMS could ask the student to imagine themselves standing on a clock face and turn towards their ten o'clock. If the student can picture an analog clock face, this method easily communicates the precise angle of rotation needed to get from one area to the next. We use these insights to inform our implementation of Clew3D, an iOS-based prototype built on top of Clew. While there are many systems that provide automatic guidance along a route (*e.g.*, [1, 5, 6, 10]) the two novel contributions of our work are to (1) co-design with COMS to leverage traditional techniques for route narration to create an automated route narration system and (2) investigate the extent to which a smartphone-based LIDAR system can automatically infer relevant environmental information to generate a useful narration of a route. By comparing a narration generated by Clew3D to a narration of the same route provided by a COMS, we seek to analyze the limitations and capabilities of automatic narrations and also to guide future work in this field.

2 O&M TECHNIQUES AND TERMINOLOGY

O&M specialists are trained professionals who work with individuals who are BVI to help them travel independently. Though the pedagogy of teaching a route is tailored to each individual student, there are certain techniques and terminology that O&M specialists are trained to use [8]. Some common techniques, reported by the O&M specialists during co-design of the proposed system include:

- **Shorelining**: Following a "shoreline" formed by the intersection of two surfaces, usually with a cane [9].
- **Squaring off**: Instructing the student to orient themselves with their back against a wall (*e.g.*, prior to navigating across an open space).
- **Clock or compass directions to indicate angle of rotation**: Instructing the student to angle themselves as if they were standing on top of a clock face, with 12 o'clock pointing straight ahead.
- **Internalize shape of the route**: Helping to develop a mental intuition for a route by describing its shape, often as a letter, *i.e.* "L" or "C".
- **Split route into distinct sections**: Breaking down a route into small sections to provide context along the route and make it easier to internalize.
- **Highlighting key features**: Landmarks (any familiar object or stimulus that is easily recognized and has a known, constant location in the field) provide reliable information about the position and line of direction (*e.g.*, door, stairwell, lamppost, hazard or directional tiles). Additionally, O&M instructors may also provide "clues", which are similar to landmarks except they are sensory, and are not necessarily physical objects (*e.g.*, the sound of a busy hallway or smell of a coffee shop) [4].

Given these core techniques for teaching a route, a COMS works with a student to formulate a set of instructions along with relevant contextual information (*e.g.*, the location of landmarks) to help the student travel the route. The precise set of techniques that are used for any given route may vary depending on particular route characteristics and what resonates with the student. In the current version of Clew3D, we consider shorelining, clock directions, splitting the route, and highlighting key features. Other techniques will be implemented as future work.

3 METHODS

3.1 Recording Data, Plane Detection

Based on the O&M techniques described previously, we built a system that allows for the detection of landmarks and features using an iPhone equipped with a LIDAR and camera. Clew3D leverages Apple's ARKit library [7], which uses the LIDAR to detect the location, extent, and orientation of surfaces. This gives Clew3D improved spatial recognition and surface detection, as the scanner can quickly build accurate depth maps, allowing for more complex computer vision analysis. ARKit also classifies surfaces that it encounters as one of seven built-in classification types: floor, wall, ceiling, door, table, seat, or unknown. We added our own stair-detection algorithm, since O&M specialists reported stairwells to be a key landmark. Clew3D tracks the phone's position in space throughout the recording phase, capturing internal keypoints along a route as well as any detected surfaces from the LIDAR.

3.2 Generating a Narration

To generate a route narration, we extract spatial relationships between detected surfaces and the trajectory of the route. We defined algorithms to detect potential shorelines (a long plane running parallel to the direction of travel), hallways (long, parallel planes on both sides of the segment), doors, and stairs. If any of these cases are detected for a given segment, we add a description to the segment's narration. We also check the distance between the current and subsequent waypoint, and the angle between the current segment and the previous segment. The angle is presented as either a clock-face direction, or simply "right", "left" or "straight ahead". This process is repeated for each segment of the route, resulting in a series of sentences that can be presented to an individual who is BVI. Decisions about the wording and types of information provided in the generated narration are informed by the aforementioned O&M techniques.

3.3 Narration Comparison

We present two narrations of a route beginning in a parking lot and ending in an apartment bedroom. The first was created by a COMS, and the second was generated by Clew3D. When comparing the two, it is evident that while Clew3D succeeds in conveying the general structure of the route, it misses important landmarks including the single step at the building entrance, the hedge, and the shoreline path from the hedge to the building.

3.3.1 *Narration by an O&M Instructor.*

(1) With the parking lot on the left continue straight to locate a hedge
(2) Turn right keeping the hedge on your left
(3) At the end of the hedge turn left
(4) Shoreline path on right side until you reach a building on your left
(5) Move to the left side and shoreline the building

(6) At the end of the building do a sharp left turn to locate single step to the front door which will be on your left
(7) Continue straight as your enter into the living room passing a TV cabinet with the kitchen on your right
(8) Proceed up first flight of stairs
(9) Turn right at landing and continue up second flight
(10) At top of stairs continue straight approximately 3 metres to locate the bedroom door (END)

3.3.2 *Clew3D's Automatic Narration.*

(1) In an open area, turn to face your 2 o'clock and walk 12 feet
(2) Turn to face your 11 o'clock and walk 2 feet
(3) Walk 5 feet forwards
(4) In an open area, walk 2 feet forwards
(5) Turn to face your 11 o'clock and walk 12 feet, shoreline the wall to your left side
(6) Turn to face your 10 o'clock and walk 10 inches
(7) Turn to face your 11 o'clock and go through the door 2 feet ahead
(8) Turn right and walk 5 feet
(9) Turn right and proceed upstairs, shoreline the wall to your left side
(10) Turn to face your 1 o'clock and walk 11 inches
(11) Go through the door 2 feet ahead (END)

4 CONCLUSION AND FUTURE WORK

Being able to detect all potential O&M landmarks including 1) strategic landmarks based on their position along the route, such as floor changes or objects near turning points, 2) contextual landmarks that could be points of interest, such as the kitchen, trash can, or bathroom, and 3) text in an environment, such as exit signs, is key to a robust implementation. Additionally, individuals who are BVI and expert travellers may find technical suggestions such as shorelining to be too basic. Future work could allow for the user to choose the information they want in their generated narration, e.g., allowing highly skilled travellers to receive only contextual information regarding their environment. Collaborating with both individuals who are BVI and COMS on the content, choice of words, and presentation of route narrations generated by Clew3D will thus be an important step moving forward.

ACKNOWLEDGMENTS

We would like to thank the individuals who provided us with their insights, including Laura Bozeman, PhD, affiliated with University of Massachusetts Boston and Cristian Bernareggi, affiliated with Universitá degli Studi di Milano. This work was supported by the National Science Foundation under Grant No. CISE-2007824.

REFERENCES

[1] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: A Navigational Cognitive Assistant for the Blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 90–99. https://doi.org/10.1145/2935334.2935361
[2] Dragan Ahmetovic, João Guerreiro, Eshed Ohn-Bar, Kris M Kitani, and Chieko Asakawa. 2019. Impact of Expertise on Interaction Preferences for Navigation Assistance of Visually Impaired Individuals. In *International Cross-Disciplinary Conference on Web Accessibility (W4A)*.
[3] Dragan Ahmetovic, Sergio Mascetti, Cristian Bernareggi, João Guerreiro, Uran Oh, and Chieko Asakawa. 2019. Deep Learning Compensation of Rotation Errors During Navigation Assistance for People with Visual Impairments or Blindness. *ACM Transactions on Accessible Computing (TACCESS)* 12, 4 (2019), 1–19.
[4] Diane L. Fazzi, Janet M. Barlow, Everett Hill, and Purvis Ponder. 2017. *Orientation and mobility techniques: a guide for the practitioner* (second edition ed.). AFB Press, American Foundation for the Blind.
[5] Aura Ganz, James M Schafer, Yang Tao, Carole Wilson, and Meg Robertson. 2014. PERCEPT-II: Smartphone based indoor navigation system for the blind. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 3662–3665.
[6] Else M Havik, Aart C Kooijman, and Frank JJM Steyvers. 2011. The effectiveness of verbal information provided by electronic travel aids for visually impaired persons. *Journal of Visual Impairment & Blindness* 105, 10 (2011), 624–637.
[7] Apple Inc. 2018. ARKit Apple Developer. https://developer.apple.com/arkit/.
[8] William Henry Jacobson. 1993. *The art and science of teaching orientation and mobility to persons with visual impairments.* American Foundation for the Blind.
[9] Hamish Maclennan. 2015. A MEANINGFUL ASSESSMENT PROCESS FOR WAYFINDING PATH SHORELINES. (05 2015). https://doi.org/10.13140/RG.2.1.1739.1529
[10] Elke Mattheiss and Elmar Krajnc. 2013. Route descriptions in advance and turn-by-turn instructions-usability evaluation of a navigational system for visually impaired and blind people in public transport. In *International Conference on Human Factors in Computing and Informatics*. Springer, 284–295.
[11] William R Wiener, Richard L Welsh, and Bruce B Blasch. 2010. *Foundations of orientation and mobility*. Vol. 1. American Foundation for the Blind.
[12] Chris Yoon, Ryan Louie, Jeremy Ryan, MinhKhang Vu, Hyegi Bang, William Derksen, and Paul Ruvolo. 2019. Leveraging augmented reality to create apps for people with visual disabilities: A case study in indoor navigation. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 210–221.