Range-IT: Detection and Multimodal Presentation of Indoor Objects for Visually Impaired People

Limin Zeng
Gerhard Weber
HCI, TU Dresden
Dresden 01187, Germany
Limin.zeng@tu-dresden.de
Gerhard.weber@tu-dresden.de

Markus Simros
Vodafone Chair, TU Dresden
Dresden 01187, Germany
Markus.simros@tu-dresden.de

Peter Conradie
Jelle Saldien
Industrial Systems Engineering and Product Design, Ghent University, Ghent, Belgium
Peter.conradie@ugent.be
Jelle.saldien@ugent.be

Ilse Ravyse
SoftKinetic
Brussel, Belgium
Ilse.Ravyse@sony.com

Jan van Erp
Tina Mioch
TNO
Soesterberg, the Netherlands
jan.vanerp@tno.nl
tina.mioch@tno.nl

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).
MobileHCI '17, September 4–7, 2017, Vienna, Austria
ACM 978-1-4503-5075-4/17/09.
https://doi.org/10.1145/3098279.3125442

Abstract
In the paper we present our Range-IT prototype, which is a 3D depth camera based electronic travel aid (ETA) to assist visually impaired people in finding out detailed information of surrounding objects. In addition to detecting indoor obstacles and identifying several objects of interest (e.g., walls, open doors and stairs) up to 7 meters, the Range-IT system employs a multimodal audio-vibrotactile user interface to present this spatial information.

Author Keywords
Visual impairments; electronic travel aids; vibrotactile; multimodal; spatial information; cognitive maps.

ACM Classification Keywords
H.5.2. [User Interfaces] Haptic I/O, Auditory feedback;

Introduction
Mobility can be a challenge for people with visual impairments, especially in unfamiliar places. While traditional mobility aids such as white canes and guide dogs provide valuable assistance, they have certain limitations. For example, a white cane helps to detect obstacles on the ground within the range of the cane (Ca. 1m), but fails to detect hanging obstacles [2, 6];

and a guide dog always needs a long-term training and can be costly to maintain.

Since the 1960's, there has been a large number of electronic aids proposed (ETAs), with the aim of overcoming the shortcomings of traditional mobility aids. Systems typically employ different types of sensors, like ultrasonic sensors [5], laser sensors [1] and camera sensors [3, 7], and auditory and/or haptic user interfaces. Vibrotactile feedback [8] provides a simple tactile representation of approaching obstacles, and the pin-matrix display based tactile representation offers users a spatial layout of surrounding multiple objects [7], but it is time- and cognition-consuming to explore the tactile display.

Considering the recent achievements in computer vision and machine learning, it is possible that camera-based systems (e.g., RGB or depth cameras) enable reliably identification of daily objects. Once those technologies are adopted in modern ETAs, it will help visually impaired people not only detect and avoid obstacles in a larger range, but also accurately recognize objects and explore the surroundings independently.

In this paper, we present our novel ETA prototype, namely Range-IT, developed through a user centered approach [4]. The Range-IT system is a 3D depth camera based modern ETA to support visually impaired people to detect obstacles within 7 meters, and identify objects of interest (OOIs), such as walls, open doors and stairs. An audio-vibrotactile user interface is used to represent simple obstacle information (e.g., distance and direction) while walking and detailed obstacle information (including object category) while standing still and exploring the surroundings. A smart IMU-based trigger is proposed to help users interact with Range-IT system conveniently, while considering users' movement, like walking/stopping and turning/scanning. The Range-IT system is designed to be used together with a common white cane, for instance to detect obstacles in the close blind area which are out of the field of view (FoV) of the camera.

Below we will further discuss the technical details of the system.

The Range-IT System

The Range-IT system is a camera-based mobility aid to enhance white cane users' capabilities of perception of surrounding obstacles/objects while travelling. In addition to a 3D depth camera, there are several components in the Range-IT system, see Figure 1.

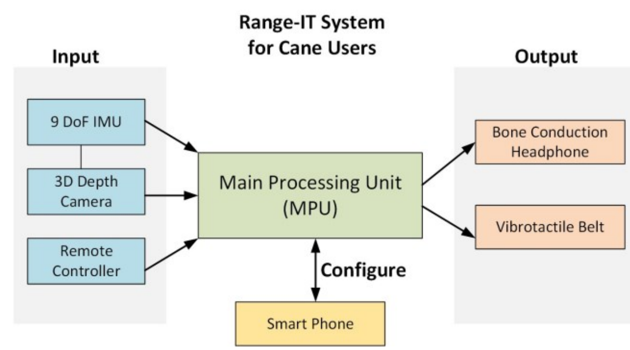

Figure 1: The system component of the Range-IT system

System Component

Figure 2 illustrates the main hardware components of the mobile Range-IT system, which are:

- **3D depth camera**: a light 3D Time-of-Flight (ToF) depth camera is provided by SoftKinetic, which includes a long work range (up to 7m) 3D depth camera (Weight: 200 grams; Dimension: 15 x 3 x 3 cm; Resolution 320 x 240; frame rate: 25 – 60 fps; horizontal field of view (FoV): 74°, vertical FoV: 87°), and a RGB camera The next generation system can deploy a compact DepthSense camera having smaller size and lower power consumption.
- **9 DoF IMU**: a Sparkfun 9 DoF Razor IMU is placed on the top of the 3D camera, in order to detect the tilt angle of the camera and the users' movement (walking/stopping, steps and turning);
- **Remote controller**: a wired remote controller with 3 keys is made by 3D printing, and the 3 keys are used to trigger Walking Mode, Exploration Mode, and mute the audio output when needed;
- **Main processing unit** (MPU): a credit card size embedded board, Odroid-U3, which consists of a 1.7GHz Quad-Core CPU and 2 GB RAM, is to do the main calculation and control processing;
- **Bone conduction headphone**: allows users to hear environmental sounds while using the system;
- **Vibrotactile belt**: is placed around user's waist and consists of 7 vibrators, to inform object's' direction and distance information;
- **Smart phone**: to configure the Range-IT system according to users' needs, such as the vibration intensity, audio files for alert, language, and the battery status;

Figure 2: The system overview of the Range-IT system

3D Scene Processing

The 3D scene processing consists of 4 main steps: pre-processing (e.g., removing noise points, and downsampling), plane fitting to identify floors, object segmentation and calculation of their spatial information (i.e., type (grounded or hanging), distance, direction and size), and object recognition. With the help of a Euclidean Cluster Extraction in Point Cloud Library[1], the Range-IT system is able to detect an objects in a minimax size of 30 x 25 cm within the work range. A SVM[2] (Support Vector Machine) based approach is used to recognize objects of interest (e.g., walls, open doors and ascending stairs). To recognize

[1] Point Cloud Library: http://pointclouds.org

[2] Lhttps://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf

more objects, a deep learning based method will be adopted in the coming future.

Working Modes

The Range-IT system has three working modes, which are:

- **Walking Mode**: when a Range-IT user is walking, the system only classifies obstacles into 2 categories (i.e., grounded obstacles and hanging obstacles), and their attributes (i.e., distance and direction to the user, and size). In Walking Mode, the system will automatically and continuously process the latest 3D image frame, and inform the obstacles' distance and direction via the tactile belt. Specifically, for hanging obstacles, a sonification sound will be played.
- **Exploration Mode**: when a user is standing and triggers the Exploration Mode by pressing the remote controller, the system will detect objects in FoV and recognize objects of interest (e.g., walls, open doors and stairs), and classify the rest objects into 2 categories (i.e., grounded obstacles and hanging obstacles). The objects' detailed information is described verbally.
- **Configuration Mode**: a user can configure the system in a personalized way, such as the vibration intensity, audio files for alert, language, and the battery status. The system uses common and basic vocabulary (e.g., use the word "unusual" rather than the word "arcane").

The Concept of Corridor Model

In Walking Mode, there are 3 corridors, which separate the area before the camera, see Figure 3. The width of each corridor is 1 m. In particular, the middle corridor is the most important one, where users walk. Once there are one or more obstacles in the 3 corridors, the 3 corresponding tactors will vibrate in different patterns, depending on the distance of the obstacles (i.e., near or far).

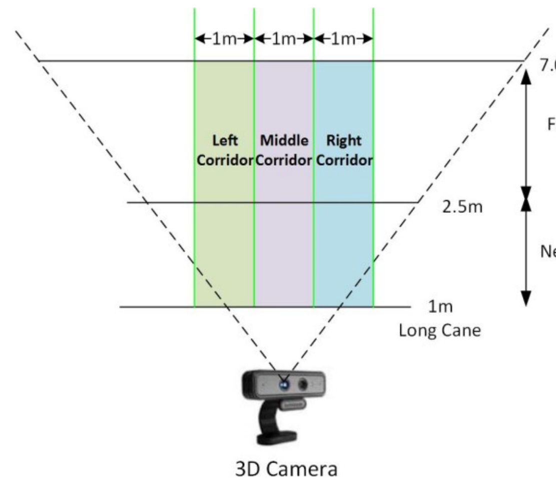

Figure 3: The concept of corridor model in Walking Mode

User Interfaces

- **Smart Movement Trigger Input**: It is time- and energy-consuming to process each frame, specifically for a mobile ETA. In the Range-IT system, a smart movement trigger is proposed to save energy and response users' movement quickly. The system would process the latest 3D image once the movement trigger is activated when a considerable number of steps (e.g., 2-3 steps) or a turning angle (e.g., 45°) is detected by the 9 DoF IMU sensor.

  The IMU sensor is fixed on the top of the camera, see Figure 4.

Figure 4: The 9 DoF IMU fixed on the top of the 3D camera

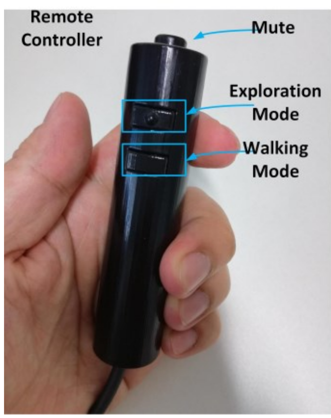

Figure 5: The wired remote controller with 3 keys

- **Key-based Controller Input**: In addition to the smart movement trigger, users are able to start Walking Mode and Exploration Mode by pressing the keys on the remote controller, as well as muting the system audio and/or tactile outputs, see Figure 5.
- **Vibrotactile Belt Output**: In total, 7 tactors are placed in the tactile belt, see Figure 6. In Walking Mode, the Tactor L3, M and R3 will be used to indicate the obstacles in the left, middle and right corridor, respectively. To help precisely perceive the direction of objects in Exploration Mode, the 5 tactors (i.e., L2: -60°, L1: -30°, M: 0°, R1: 30°, and R2: 60°) are used. The vibrotactile messages will present obstacles/objects' distance with 2 categories: near (distance < 2.5 m or far (distance > 2.5m) and direction (i.e., 10 o'clock: -60°, 11 o'clock: -30°, 12 o'clock: 0°, 1 o'clock: 30°, and 2 o'clock: 60°).
- **Auditory Output**: The auditory output consists of two types of sounds: sonification sounds and verbal sounds. In addition to describing the system errors or reminders, verbal sounds are used in Exploration Mode to describe detailed information of objects, for instance, "Hanging at 4 meters, 12 o'clock". During the development of the system, users pointed out they did not prefer a number of auditory signals while walking except a specific sound for indicating hanging objects. Therefore, a special sonification sound for hanging objects is imported in the final system.

To reduce users' workload in Walking Mode and assist them to acquire detailed spatial information of surrounding objects in Exploration Mode, the Range-IT system implement an audio-vibrotactile user interface by integrating the vibrotactile output and the auditory output.

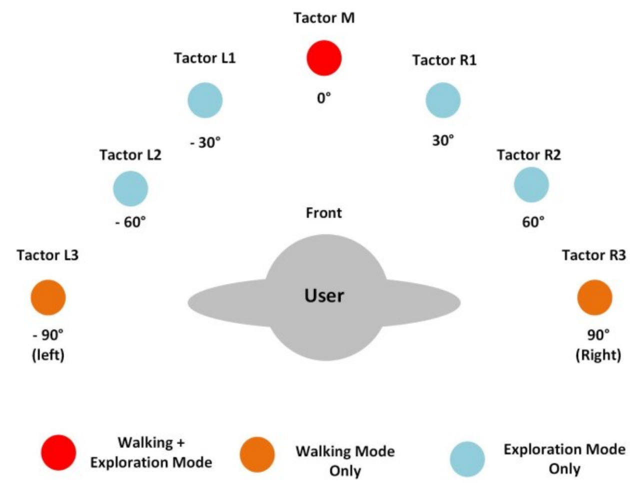

Figure 6: The layout of the tactors on the tactile belt

Acknowledgements

We thank all the volunteers, publications support, staff, and authors who wrote and provided helpful comments on previous versions of this document. The authors gratefully acknowledge the grant from EU FP7 (#605998).

References

1. Malvern J. Benjamin and Nazir A. Ali. 1974. An improved laser cane for the blind. In *Proc. SPIE 0041, Developments in Laser Technology II*, 107 (March 1, 1974); doi:10.1117/12.953841.

2. Roberto Manduchi and Sri Kurniawan. 2011. Mobility-related accidents experienced by people with visual impairment. *Research and Practice in Visual Impairment and Blindness*, 4(2), 44-54.
3. Peter B. L. Meijer. 1992. An experimental system for auditory image representations. *IEEE Trans. on Biomedical Engineering*, 39 (1992), 112-121.
4. Peter. Conradie, De Marez, and Jelle Saldien. 2015. Participation is Blind: Involving Low Vision Lead Users in Product Development. 6th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Infoexclusion (DSAI 2015) (Bonn, 2015).
5. Iwan Ulrich and Johann Borenstein. 2001. The GuideCane-applying mobile robot technologies to assist the visually impaired. *IEEE Trans. on SMC*, 31 (2001), 131-136.
6. Limin Zeng. 2015. A Survey: Outdoor Mobility Experiences by the Visually Impaired. In *Proc. Mensch & Computer Workshopband 2015*: 391-397
7. Limin Zeng, Denise Prescher, and Gerhard Weber. 2012 Exploration and avoidance of surrounding obstacles for the visually impaired. In *Proc. ACM ASSETS 2012*, 111-118.
8. Michael Zöllner, Stephan Huber, Hans-Christian Jetter, and Harald Reiterer. 2011. NAVI - A proof-of-concept of a mobile navigational aid for visually impaired based on the Microsoft Kinect. In *Proc. INTERACT 2011*, 584-587.