NAVI – A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect

Michael Zöllner, Stephan Huber, Hans-Christian Jetter, Harald Reiterer

University of Konstanz, Germany
{michael.zoellner, stephan.huber, hans-christian.jetter, harald.reiterer}@uni-konstanz.de

Abstract. We present a proof-of-concept of a mobile navigational aid that uses the Microsoft Kinect and optical marker tracking to help visually impaired people find their way inside buildings. The system is the result of a student project and is entirely based on low-cost hard- and software. It provides continuous vibrotactile feedback on the person's waist, to give an impression of the environment and to warn about obstacles. Furthermore, optical markers can be used to tag points of interest within the building to enable synthesized voice instructions for point-to-point navigation.

Keywords: Vibrotactile User Interface, Navigation User Interface.

1   Introduction

Wayfinding is a cognitive demanding, error-prone task for visually impaired persons. It is described by Long and Hill as "the process of navigating through an environment and travelling to places by relatively direct paths" [8]. As such, it depends heavily on both "sensing of the immediate environment for obstacles and hazards" [9] and "navigating to remote destinations beyond the immediate perceptible environment" [9]. These two aspects of wayfinding are also called *micro-navigation* and *macro-navigation* [10]. Although both aspects are part of successful wayfinding, there are
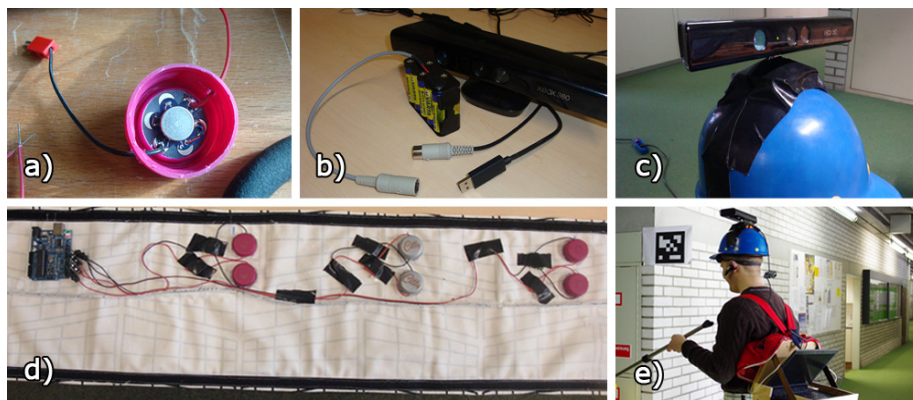

Figure 1: (a) LilyPad vibe board in a plastic cap. (b) Mobile Kinect camera with battery pack. (c) Kinect helmet. (d) Vibrotactile Waist Belt. (e) Complete setup with backpack.

only few systems that try to integrate them together (e.g. [1, 7]). These systems typically try to substitute their users' limited or missing visual capabilities by giving helpful feedback via different sensory channels. Sound and speech are probably the most frequently used sensory channels in this domain. However, as environmental auditory signals are very important for persons with limited sight, acoustic feedback must be used carefully and is not appropriate in all cases. Especially when it is used for continuous feedback - which is very important in micro-navigation scenarios such as obstacle detection - acoustic feedback can become annoying and distract users from the natural aural environment. In macro-navigation scenarios, speech output - similar to that of car navigation systems – seems to be a reasonable choice. However, as different users interpret these signals differently, it is not easy to give unambiguous navigation instructions. Even worse, workload may differ if navigation instructions are given by sighted persons instead of visually impaired persons [2]. It "seems to be a hopeless situation for the ability of language to convey spatial information" [6]. An alternative is vibrotactile feedback that is used for navigation purposes in various ways, ranging from vibrating headbands [4] to shoulder tapping systems [12]. As vibrotactile signals are less disruptive than sound or speech, they are an appropriate alternative to give continuous feedback. Thus, our system employs a hybrid approach that augments synthesized speech instructions for macro-navigation with vibrotactile feedback for micro-navigation using only low-cost and self-made technology.

2 System Description

The initial design of our proof-of-concept was created during HCI coursework at our university. Our design goal was to build a system that enables and/or facilitates both micro-navigation and macro-navigation. For this we utilized the Microsoft Kinect sensor, a vibrotactile waist belt built with Arduino LilyPad[1] vibe boards and a simple backpack construction that carried the laptop and enabled quick debugging (Fig.1 e).

For detecting the immediate surroundings, we reversed the standard operating principle of the Kinect. Instead of a static Kinect that tracks moving objects, we track the static environment with a moving head-mounted Kinect (Fig.1c). To power the mobile Kinect we use a 12V battery pack that lasted for about 5 hours during our tests (Fig.1 b). The vibrotactile output is provided by a waist belt that contains three pairs of Arduino LilyPad vibe boards (Fig.1a). These are fixed into plastic bottle caps to amplify the perceived vibration (Fig.1d). The speech output is provided by an ordinary Bluetooth headset for mobile phones (Fig.1e).

Micro Navigation – In contrast to Lee et al., who use vibrotactile signals to encode navigation instructions [7], we use them to inform users of persons or other obstacles in their way, thereby enabling secure micro-navigation. Erp et al. showed that encoding distances on a vibrotactile display is realized best with temporal patterns instead of changes in frequency or amplitude [5]. Based on several informal evaluations with different distance signals, we decided for a simple yet robust distance encoding with just three different outputs: no output meaning no obstacle, a
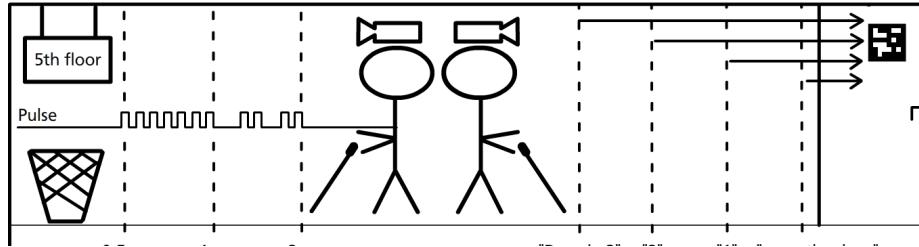

Figure 2. Vibrotactile cues and navigation instructions depend on the distance to the obstacle (left side) or marker (right side).

500ms pulsed vibration for signaling obstacles between 1-2m distance and a continuous vibration for obstacles closer than 1m (Fig. 2). Erp et al. also investigated that spatial information is encoded most effectively by applying the vibrotactile signals to different body locations [5]. Thus we located the vibe boards on the left, right and center of the person's waist to indicate the direction in which an obstacle was detected. Unlike Cardin et al. [3], who detect only the closest obstacle, we are able to detect the closest obstacles in the left, right and center region of the Kinect's "field of view". The signal processing software and the control of the vibe boards is implemented with C#/.NET and uses a managed wrapper of the OpenNI framework[2]. Obstacles in each of the image's regions are identified via a depth histogram. We assume that nearby and large objects are potentially harmful to the person. Thus, the obstacle detection algorithm moves a depth window of 120mm size from left to right (respectively near to far) over the histogram and stops, if the pixel area of that depth window exceeds a certain threshold area (approx. 4% of a region). The average depth value of the current depth window is then mapped to the pulse of the vibe board.

The overall minimalistic design of the vibrotactile output (3 output channels with 3 different signals each) is based on several iterations with informal experiments, where users felt more comfortable and performed better with less sensory information. We therefore chose to minimize the output's complexity thereby also minimizing the users' cognitive load.

Macro-Navigation – We use synthesized voice as feedback mechanism for macro-navigation. Since navigation instructions are very short (e.g. "turn left") and don't occur continuously, they only slightly interfere with environmental sounds. Our low-cost solution uses printed augmented reality markers (Fig.1e) that can be used to tag points of interest on the desired route. We detect the markers via Kinect's RGB camera using a managed wrapper of the ARToolkitPlus[3]. By integrating depth information of the Kinect, we facilitate different navigation instructions based on the person's distance to the marker. For example, when walking towards a door, the instructions will be as follows: "Door in 3", "2", "1", "Open the door" (Fig. 2). Our approach has the drawback that every route has to be signposted individually. While the effort to create and maintain such a relative positioning system is higher compared to a map-based absolute positioning system like GPS, it has the essential advantage of being available inside buildings.

---

[2] https://github.com/kobush/ManagedOpenNI (Retrieved June 2, 2011)
[3] http://code.google.com/p/comp134artd (Retrieved June 2, 2011)

Future Work

In this work we presented a system that facilitates micro- and macro-navigation for visually impaired persons. We plan to iteratively improve this proof-of-concept in the future: (1) We observed that users don't walk straightforward and frequently don't approach markers from the front but from varying angles. This rendered some navigation instructions useless, as they were designed for frontal approaches. Future work should adapt navigation instructions to the angle from which the user is approaching. (2) Like Ram and Sharf [11] we would like to discriminate between persons and inanimate objects. This would help to reduce the amount of system output while moving in a crowded environment and could prevent users from sensory and cognitive overload. (3) Finally, we plan to conduct a formative user study with visually impaired users to evaluate navigation instructions and vibrotactile signals. A further summative study could then reveal how NAVI compares to the conventional white cane in terms of navigation effectiveness, efficiency and cognitive load.

References

1. e-adept: Electronic assistance for disabled and elderly pedestrians and travelers. Retrieved June 2, 2011, from http://www.eadept.se.
2. N. A. Bradley and M. D. Dunlop. An experimental investigation into wayfinding directions for visually impaired people. *Personal Ubiquitous Computing*, 9:395–403, November 2005.
3. S. Cardin, D. Thalmann, and F. Vexo. A wearable system for mobility improvement of visually impaired people. *Vis. Comput.*, 23:109–118, January 2007.
4. A. Cassinelli, C. Reynolds, and M. Ishikawa. Augmenting spatial awareness with haptic radar. In *Proc. ISWC '06*, pages 61–64, 2006.
5. J. Erp, H. Veen, C. Jansen, and T. Dobbins. Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception*, 2(2):106–117, 2005.
6. N. Franklin. Language as a means of constructing and conveying cognitive maps. In Juval Portugali, editor, *The Construction of Cognitive Maps*, volume 32 of *GeoJournal Library*, pages 275–295. Springer Netherlands, 1996.
7. J. Lee, E. Choi, S. Lim, and B. Shin. Wearable computer system reflecting spatial context. In *Proc. IWSCA '08*, pages 153–159, Washington, DC, 2008. IEEE Computer Society.
8. R. G. Long and E. W. Hill. Establishing and maintaining orientation for mobility. In W. R. Wiener B. B. Blasch and R. L. Welsh, editors, *Foundations of Orientation and Mobility*, chapter 2, pages 39–59. AFB Press, second edition, 1997.
9. J. M. Loomis, R. G. Golledge, and R. L. Klatzky. Gps-based navigation systems for the visually impaired. In W. Barfield and T. Caudell, editors, *Fundamentals of Wearable Computers and Augumented Reality*. L. Erlbaum Associates Inc., Hillsdale, NJ, 2000.
10. H. Petrie. User requirements for a gps-based travel aid for blind people. In *Proc. of the Conference on Orientation and Navigation Systems for Blind Persons*. RNIB, 1995.
11. S. Ram and J. Sharf. The people sensor: A mobility aid for the visually impaired. In *Proc. ISWC '98*, pages 166–, Washington, DC, 1998. IEEE Computer Society.
12. D. A. Ross and B. B. Blasch. Wearable interfaces for orientation and wayfinding. In *Proc. ASSETS '00*, pages 193–200, New York, NY, 2000. ACM.

NAVI - A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect

Michael Zöllner, Stephan Huber, Hans-Christian Jetter, Harald Reiterer

Abstract

We present a proof-of-concept of a mobile navigational aid that uses the Microsoft Kinect and optical marker tracking to help visually impaired users find their way inside buildings.

The system is the result of a student project and is entirely based on low-cost hard- and software it provides continuous vibrotactile feedback on the person's waist, to give an impression of the environment and to warn about obstacles.

Furthermore, optical markers can be used to tag points of interest within the building to enable synthesized voice instructions for point-to-point navigation.

1 The head-mounted Kinect scans the immediate environment of the visually impaired person and generates depth and RGB images for further processing.

2 We mobilized the kinect camera by connecting a 12V battery pack instead of a standard power supply.
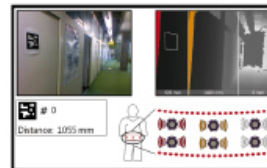
3 The depth information of the Kinect is mapped by our software onto three pairs of Arduino LilyPad vib boards located at the left, center and right of the waist.

We pinned several optical markers on the walls and doors of our building to model a certain route from one room to another.

4

5 The debugger UI gives a quick overview of the system's state

Human-Computer Interaction Group
University of Konstanz
http://hci.uni-konstanz.de/

Contact:
Michael Zöllner
Tel. +49 7531 88-3534
Fax +49 7531 88-4772
michael.zoellner@uni-konstanz.de

Stephan Huber
Tel. +49 7531 88-2868
Fax +49 7531 88-4772
stephan.huber@uni-konstanz.de