BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians

Conference Paper · May 2019

DOI: 10.1145/3290605.3300282

7 authors, including:


Seita Kayukawa
IBM Research
18 PUBLICATIONS **238** CITATIONS
SEE PROFILE


Keita Higuchi
The University of Tokyo
28 PUBLICATIONS **756** CITATIONS
SEE PROFILE


João Guerreiro
University of Lisbon
72 PUBLICATIONS **1,217** CITATIONS
SEE PROFILE


Shigeo Morishima
Waseda University
368 PUBLICATIONS **3,949** CITATIONS
SEE PROFILE

BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians

Seita Kayukawa
Waseda University
Carnegie Mellon University
k940805k@ruri.waseda.jp

Keita Higuchi
University of Tokyo
khiguchi@iis.u-tokyo.ac.jp

João Guerreiro
Carnegie Mellon University
jpvguerreiro@cmu.edu

Shigeo Morishima
Waseda Research Institute for Science and Engineering
shigeo@waseda.jp

Yoichi Sato
University of Tokyo
ysato@iis.u-tokyo.ac.jp

Kris Kitani
Carnegie Mellon University
kkitani@cs.cmu.edu

Chieko Asakawa
IBM Research
Carnegie Mellon University
chiekoa@cs.cmu.edu

ABSTRACT

We present an assistive suitcase system, BBeep, for supporting blind people when walking through crowded environments. BBeep uses pre-emptive sound notifications to help clear a path by alerting both the user and nearby pedestrians about the potential risk of collision. BBeep triggers notifications by tracking pedestrians, predicting their future position in real-time, and provides sound notifications only when it anticipates a future collision. We investigate how different types and timings of sound affect nearby pedestrian behavior. In our experiments, we found that sound emission timing has a significant impact on nearby pedestrian trajectories when compared to different sound types. Based on these findings, we performed a real-world user study at an international airport, where blind participants navigated with the suitcase in crowded areas. We observed that the proposed system significantly reduces the number of imminent collisions.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility technologies**; • **Social and professional topics** → *People with disabilities*.

KEYWORDS

Visual impairments; obstacle avoidance; collision prediction; pedestrian detection; blind navigation; path clearing.

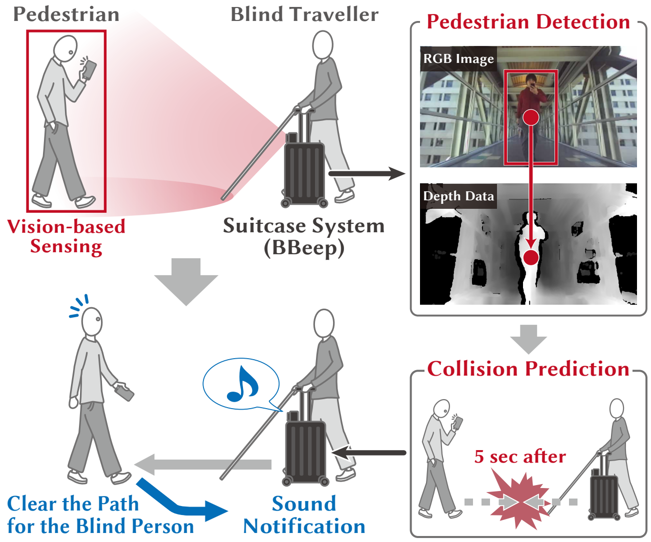

Figure 1: BBeep is an assistive suitcase system that uses sonic feedback to alert both the blind user and nearby sighted pedestrians about potential risks of collision.

ACM Reference Format:
Seita Kayukawa, Keita Higuchi, João Guerreiro, Shigeo Morishima, Yoichi Sato, Kris Kitani, and Chieko Asakawa. 2019. BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300282

1 INTRODUCTION

Blind people face significant challenges when navigating public spaces due to the lack of visual sensing. Recent research using computer vision aimed to assist blind users' orientation and mobility skills for avoiding potential obstacles or hazards [7, 16, 22, 23, 29, 37, 46, 49, 50]. While these systems are often able to detect static obstacles, the detection

---

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300282

and avoidance of collisions with dynamic elements, in particular pedestrians, is still relatively unexplored in the literature. Technical challenges aside, one possible reason for a lack of work on dealing with dynamic elements is the assumption that sighted pedestrians are aware of blind people and therefore will always clear the path for them. However, this is not always the case as sighted people may be looking at their smartphone, talking with others, or facing another direction (looking at a board or TV). In such scenarios, blind people face significant risks of collision with other pedestrians.

We present an assistive suitcase system, BBeep, that uses a sonic collision warning system to alert both the blind user and nearby sighted pedestrians about potential risks of collision (Figure 1). This approach extends common sonic warning systems that are used to clear the path for moving vehicles, such as airport carts driving through crowded terminals or large trucks driving in reverse. This work leverages the simple fact that sighted persons can quickly get out of the way of a blind person who is walking, if they are given appropriate information about a blind person's presence. However, our work aims to go beyond the paradigm of constantly playing a sound to convey the user's presence, as the constant emission of loud alarm sounds can be social disruptive and make the blind user feel overly self-conscious. Instead, we present an adaptive sonic warning system that only emits sounds when needed. More specifically, BBeep is designed to consider the motion of nearby pedestrians, predict future collisions, and give sonic feedback only when necessary.

Although we explicitly target the navigation of blind people in airports, we believe that the form factor of a travel suitcase is also appropriate and brings several benefits in other real-world crowded environments such as train stations or shopping malls. For the blind user, a suitcase can often act as an extended sensing mechanism for identifying changes in floor texture or as a form of protection from collisions in very dense crowds. Even without any smart sensing, a suitcase can be used as an assistive device. In many cosmopolitan environments, a suitcase is a common object and does not draw unnecessary attention to the user. As a robotic sensing system, it also provides a convenient place to store and attach sensors, power and computing resources.

BBeep uses an RGBD camera to detect, track and predict the motion of nearby pedestrians. The RGB image is used to detect people using a convolutional neural network and the depth channel is used to estimate the distance to pedestrians. Averaged position estimates are used to estimate pedestrians' velocity and linear extrapolation is used to predict their future path. Depending on the proximity of the predicted path with the user, an appropriate sound is emitted by BBeep.

To investigate how to convey sonic feedback effectively, we performed an observational study where the suitcase-shaped system emits sounds of different types and timings. Results suggest that sound emission is an effective method to change the pedestrians' walking direction away from the platform, and that its timing has more impact than the sound types. Based on these findings, we designed the sonic interface of BBeep that used three stages of sound emissions to notify about potential collision risks with pedestrians.

In order to evaluate the effectiveness of BBeep for preventing collisions with pedestrians, we performed a study where six blind users walked with the suitcase in crowded areas of an airport. We observed that BBeep reduced the number of situations of imminent collision risk, when compared to only notifying the blind user. Participant feedback supported our hypothesis that BBeep is useful for collision avoidance in crowded public spaces. Based on our findings, we discuss future requirements towards a more flexible solution that is able to adapt to different scenarios and users.

2 RELATED WORK

Blind Navigation Systems

Prior research proposed various blind navigation systems [1, 2, 10, 13, 18, 31, 35, 36, 39]. Most systems guide blind people using turn-by-turn navigation, with technologies such as GPS [36, 39], RFID tags [2, 10, 15], and Bluetooth low-energy beacons [1, 11, 26]. Still, most systems do not consider dynamic environments and therefore are unaware of obstacles that were not in the environment, such as desks, chairs, and pedestrians. To overcome this limitation, it is important to assist blind users avoiding collisions with such elements.

Supportive Systems for Obstacle Avoidance

Besides guide dogs, white canes [5, 47] are the most common tool for blind people to find obstacles and avoid collisions. While very efficient, a user can detect an obstacle only after physically hitting it with the cane. This is undesirable, especially when the obstacle is a pedestrian. Researchers have developed supportive technologies that allow blind users to detect obstacles with non-contact sensing [4, 30, 43]. Systems that detect and provide information about obstacles (*e.g.* distance[22–24, 28, 30, 37], shape[7, 23, 29], or category[22, 37, 49]) to users often use laser [28], ultrasonic [24, 44], phone's speakers and microphones [45], or depth sensing[7, 16, 21–23, 37, 49]. However, it is still challenging for blind people to detect and avoid obstacles, in particular in very dense environments. In those scenarios, it is important to provide a safe path to blind users.

Some systems guide users around obstacles using sound feedback [29, 50] and/or haptic feedback from ground [46] or aerial robots [3]. Blind users can follow the system feedback to avoid static obstacles such as chairs, desks, and walls. While applicable to inanimate objects, some studies assumed that pedestrians, as dynamic obstacles, can avoid a blind user, and thus did not focus on supporting pedestrian collision avoidance [29, 46]. However, in public spaces (*e.g.*, airports), pedestrians may be unaware of blind users while using mobile devices or talking to others. Therefore, we are interested in investigating sound notification techniques to make pedestrians aware of the blind traveler. We expect that pedestrians will then clear the path for the blind user.

Sound Alert for Urgent Notifications

Beep sounds have been used as a means to notify people of urgent situations, such as in hospital intensive care units [34], nuclear power plants [33], and aviation [6]. Audio notifications can also alert drivers of an imminent collision or assist in navigation [32]. The relationship between user perception and different types of alert sounds plays a vital role in their usability. Several works have found that auditory parameters of beep sounds (*e.g.*, fundamental frequency, pulse rate, and intensity) affect perceived urgency levels [12, 20, 32], while others observed a trade-off between perceived urgency and annoyance levels of alert sounds [12, 19, 32].

As described above, emitting beep sounds is a common approach to notify users of urgent situations. In this paper, we thus use this type of sound to make pedestrians aware of a blind user. We investigate what types of beep sounds are effective for collision avoidance, and design a sound emission policy for our prototype system.

3 BBEEP: DESIGNING A PATH CLEARING SYSTEM

Our main goal is to ease the mobility of blind people in crowded environments. We argue that collisions with pedestrians can be avoided if both the blind user and sighted pedestrians are made aware of the collision risk. For this purpose, we developed BBeep, a sonic collision avoidance system that aims to clear the path for blind users.

Limitations of Notifying only the Blind User

Prior research on obstacle detection for blind people focuses on notifying the user alone about the presence of obstacles, prompting them to change their orientation [29, 46, 50]. Such approach increases the user's knowledge of the surroundings, but also comes with significant limitations in this context. First, actively changing the walking direction of users may be unsafe (taking the user through a different/unknown path); second, a group of pedestrians may block the entire route of a blind user; and finally, notifying users about all pedestrians in crowded environments may require complex feedback, which may be cognitively demanding to users. In addition, by focusing on obstacles in general these approaches do not take advantage of sighted pedestrians' ability to cooperate in collision prevention. We argue that conveying feedback only to the user may not be effective to avoid collisions with other pedestrians, in particular in crowded environments.

Sound Notifications for Users and Pedestrians

BBeep uses sound notifications due to the ability of sound to attract people's attention even when they are focused on something else. Although other modalities, such as visual stimuli (*e.g.*, Vection Field [17]), can also impact pedestrians' walking direction, it may not be as effective in several scenarios. For instance, the prevalence of smartphones significantly reduces sighted pedestrians' awareness of the surroundings, resulting in potential collisions [9]. Moreover, (groups of) people talking or looking at a different direction may not notice a blind person until their white-cane hits them. For that reason, our approach is inspired in the common use of beep sounds to notify pedestrians of urgent situations prompting them to clear the path. A few examples include carts in crowded airports or large motor vehicles driving in reverse. However, such approach also comes with significant challenges, since frequent emission of loud alarm sounds can be socially disruptive and make the user feel uncomfortable.

Collision Prediction to Reduce Sound Emissions

To enhance social acceptance of sound emissions, it is important to emit alert sounds only when absolutely necessary. Moreover, a collision prediction technique is required in order to decrease as much as possible the number of sound emissions, while maintaining its effectiveness. For this reason, BBeep relies on real-time pedestrian tracking and collision prediction to provide notifications only when there is a potential risk of collision. This is beneficial to reduce both collision risks and social disruption in public spaces.

4 BBEEP: IMPLEMENTATION

We developed a vision-sensing system for tracking the motion of pedestrians and predicting their future positions in order to generate an audible warning signal that will clear the path in front of a blind user (Figure 2). A stereo camera is attached to a suitcase to capture RGB images and collect depth data. One advantage of this setup is the ability to capture images without significant motion-induced blur and to perform the necessary computations *in situ*. The system detects pedestrians using RGB images and tracks their position using the depth data in real time. Based on these results, the system predicts the future positions of pedestrians and determines their risk of collision with a blind user. The system can then emit an audible alarm if necessary.

We note that there is prior work on pedestrian trajectory forecasting [25, 27, 48] and the aim of this work is not to advance the state of the art in this respect. Instead, our contribution is the analysis and development of effective sonic feedback mechanisms based on such predictive input. To this end, our main challenge is to develop a real-time forecasting technique with sufficient accuracy for collision prediction.

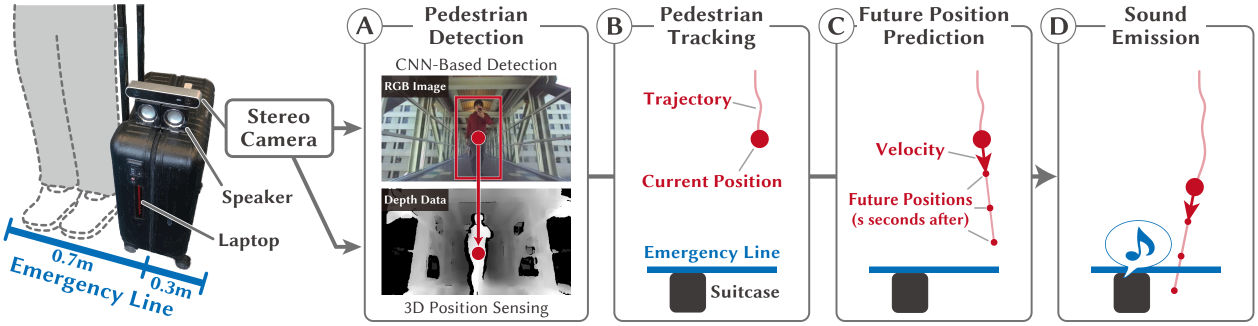

Figure 2: Overview of BBeep. The stereo camera is mounted on a suitcase and records RGB images and depth data. A) The system detects pedestrians using the RGB images and B) tracks their position using the depth data. Then, C) it predicts the future positions of each pedestrian. Finally, D) BBeep emits an audible signal if there is a risk of collision with the blind user.

Pedestrian Detection: Requirements and Design

We implemented a novel system combining stereo image sensing and a CNN-based generic object detector (YOLOv2 [40]). We use a ZED™2K Stereo Camera[1], as it has a wider horizontal field of view (90 degrees), a longer depth range (0.3 to 20 m), and a higher FPS (up to 100 Hz) than Kinect v2. The stereo camera supports a 3D odometry API that provides access to 3D movements from the camera in real time. We use this information to remove the influence of suitcase rotation.

We used YOLOv2 to detect pedestrians using the RGB streams. The method robustly detects individuals even if their body is not completely included within the camera images. We confirmed that the combination of the ZED camera and YOLOv2 can detect bounding boxes of pedestrians from a distance of 10 m. We used the central area of the bounding boxes to obtain the 3D positions of the detected pedestrians in the camera coordinate system.

Updating positions using a high FPS is important for accurately predicting pedestrians' future positions. We used a laptop computer (Intel Core i7-7700HQ CPU, NVIDIA GeForce GTX 1060 GPU) to process object detection at a rate of 15 fps, but this was insufficient to obtain satisfactory prediction accuracy. We, therefore, used a given bounding box for obtaining 3D positions, and updated it once a new detection result was available. The system thus tracks pedestrians at a frame rate in excess of 40 fps. Note that the detection and tracking processes run simultaneously on different threads.

Pedestrian Tracking

The system processes pedestrian tracking based on the detection results. We propose an algorithm to track individual pedestrians in real time. We use the following procedure to update a tracking list of detected pedestrians at each frame:

1. The tracker generates a set of bounding boxes from pedestrian detection and computes a set of 3D positions in the camera coordinate system based on the central area of the bounding boxes.
2. The tracker repeats steps (3) and (4) for each detected 3D position (the current position).
3. If there are no existing pedestrians in the list from the current position within a distance α, the tracker adds the point to the list as a new pedestrian.
4. Otherwise, the tracker updates the position of the nearest pedestrian in the list to the current position, and saves the previous position as a record of its trajectory.
5. The tracker removes pedestrians from the list if their position has not been updated by the tracker in β $frames$ of the tracking process.

Based on our observations, we set the parameter values $\alpha = 1$ m and $\beta = 5 frames$ for all of our studies.

Position Prediction and Sound Emission

The system predicts the future positions of the pedestrians in the tracking lists derived from the tracking process. The system uses the 3D positions of the pedestrians in the camera coordinate system to predict the relative speed and direction of displacement between the suitcase and each pedestrian using their current position and their trajectory. To improve the stability of the pedestrian position measurement, the system first compensates for rotations of the suitcase (camera) by rotating the detected pedestrian positions using 3D odometry information. The system then computes the expected future position $\hat{\mathbf{p}}_{t+s}$ of each pedestrian after s seconds using the $N-1$ most recent points of its trajectory as follows:

$$\mathbf{p}_\mu(i,n) = \frac{1}{n} \sum_{j=i-n+1}^{i} \mathbf{p}_j \tag{1}$$

$$\hat{\mathbf{p}}_{t+s} = \frac{\mathbf{p}_\mu(t, \frac{N}{2}) - \mathbf{p}_\mu(t-\frac{N}{2}, \frac{N}{2})}{\Delta t(\mathbf{p}_{t-\frac{N}{4}}, \mathbf{p}_{t-\frac{3N}{4}})} s + \mathbf{p}_t \tag{2}$$

---

[1] https://www.stereolabs.com/zed/

Let $\mathbf{p}_t$ be the position of a pedestrian in the camera coordinate system at time t, $\mathbf{p}_\mu(i, n)$ the average over n previous positions (from $\mathbf{p}_{i-n+1}$ to $\mathbf{p}_i$), and $\Delta t(\mathbf{p}_t, \mathbf{p}_i)$ the difference in the time stamps between $\mathbf{p}_t$ and $\mathbf{p}_i$. The system first calculates the two average positions (from $N - 1$ frames before to $N/2$ frames before, and from $N/2 - 1$ frames before to current frame) (Equation 1). The system then calculates the vector between the two average positions and predicts the future position (Equation 2). Setting $N = 32$ was found to yield stable predictions.

The system then predicts the risk of a future collision based on all the predicted positions to decide whether an alarm sound should be emitted, as outlined in Figure 2. A collision is expected when a future pedestrian trajectory crosses the "emergency line" shown in Figure 2 (D). The system computes the line connecting the current and future positions of the pedestrian as a prediction of the expected future trajectory. The system then determines the intersection between this line and the emergency line. If the intersection lies within the range of the emergency line, the risk of collision is considered significant and the system emits a warning sound. Note that this calculation does not use pedestrian height information.

5 BBEEP: DESIGNING SOUND EMISSION POLICY THROUGH AN OBSERVATION STUDY

We studied the response of pedestrians to the audible warning signals in order to design a sound-emission policy for our system. As described in Related Work, human perception of audible emergency warnings has been studied [12, 19, 32]. There is also some understanding of how a visual stimulus can cause pedestrians to redirect their trajectory [17]. Yet, little is still known about how a pedestrian reacts to an audible signal. Such insight is important for designing an effective policy of sound emissions for our path-clearing system.

We conducted an observational study in a corridor in which the suitcase-enclosed system was made to emit different types of audible alerts (beeps). We recorded pedestrians' reactions and trajectories as shown in Figure 3. We designed a set of sound-emission patterns comprising various sound types and a range of timings. The system tracked pedestrians in the corridor and emitted alerts using these patterns. We analyzed the pedestrians' trajectories to determine which sound patterns were most effective at clearing the path in front of the suitcase. Based on the outcome, we then designed our sound-emission policy for evaluation in an airport.

Sound-Emission Patterns

We designed 7 sound patterns (S1–S7), one of which (S7) was mute to serve as a baseline. The 6 non-baseline patterns featured combinations of 3 different sound types and 2 types of emission timings.

Sound	Sound type				Timing
Pattern	UL	BF	PD	IPI	s
S1	High	1000 Hz	0.1 s	0.1 s	5.0 s
S2					2.5 s
S3	Middle	400 Hz	0.1 s	0.1 s	5.0 s
S4					2.5 s
S5	Low	400 Hz	0.5 s	0.5 s	5.0 s
S6					2.5 s
S7	Without sound emission				N/A

Table 1: Sound-emission patterns. UL: urgency level, BF: base frequency, PD: Pulse duration, IPI: inter-pulse interval, and s: timing (*i.e.* emits a sound alert by considering the expected position of pedestrians after s seconds).

Beep-Alert Sound Types. We used alert sounds to represent 3 distinct levels of perceived urgency. The relationship between perceived urgency and sound parameters is well documented. We prepared 3 types of beep sounds with different urgency levels denoted High, Intermediate, and Low. Specifically, we varied the base frequency, the pulse rate, and the pitch, as given in Table 1. The values we used are based on recent research addressing sound urgency [38, 41].

Timing of sound emission. We also used different timings of sound emissions for each beep alerts. The system changes the timing by setting the collision detection parameter s seconds. If a system were to emit a sound alert immediately before a predicted collision with the blind user (*e.g.*, $s < 1$)), the pedestrian in question may not be able to avoid the collision. On the other hand, a sound alert emitted too long in advance (*e.g.*, $s = \infty$) may cause unnecessary disturbance and inconvenience and be effectively unproductive.

We selected the parameter values $s_1 = 5.0$ and $s_2 = 2.5$ seconds. The value of $s_1 (5.0\ s)$ represents the time needed to travel the furthest distance in the detection range (around 10 m) when the blind user and a pedestrian are approaching at a relative speed of 4 km/h. We also used $s_2 (2.5\ s)$ set to a half of s_1 to define a nearer threshold.

Data Collection and Analysis

This observational study considered the suitcase-enclosed system with 7 sound emission patterns placed in a straight corridor (Figure 3). The system tracked pedestrians and predicted their intersection with the emergency line in real time. The system also emitted sounds as specified by the adopted policy. All seven sound patterns were used in cycle.

Observations were conducted over more than four days, yielding 57 trajectories for each pattern (399 in total). The system recorded a trajectory and images of the closest pedestrian with a risk of collision for each sound pattern.

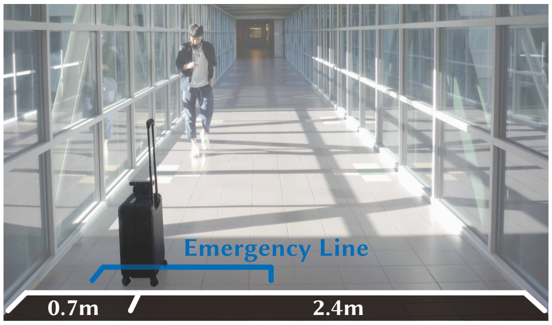

Figure 3: The suitcase-enclosed system in a corridor, equipped to emit various types and timings of beep sounds.

We analyzed the recorded datasets to identify how trajectories were affected by the emissions. Some trajectories were occasionally missing or inaccurate owing to the limitations of real-time processing, as described in Pedestrian Tracking. We therefore performed a subsequent trajectory analysis using the recorded RGB and depth images to obtain more accurate pedestrian trajectories. We used OpenPose [8], a CNN-based human-body detection software, to detect parts of pedestrians' bodies from the RGB images. We then determined the central position of the detected bodies in the depth images to obtain the 3D positions of pedestrians. We conducted the analysis for all the recorded images. The laptop used analyzed images captured at a rate of 5 fps, i.e., an insufficient rate for real-time sound notifications.

Evaluation Measurements

We measured the *"minimal distance"* between the suitcase position and a given trajectory to investigate the relevance of different sound patterns for avoiding collisions. A longer minimal distance may be interpreted as indicating that the pedestrian has avoided the blind user by a comfortable margin. These minimal distances were determined from the 3D positions returned by OpenPose.

We considered three hypotheses for the main potential factors influencing the minimal distances: the presence or absence of a sound emission (Hypothesis 1), the sound-emission timing (Hypothesis 2); and the urgency level of the emitted sound (Hypothesis 3). We tested these hypotheses using a Kruskal–Wallis test and a Mann–Whitney U test at 5% levels of significance to discern differences within sound patterns. We also saw 95 % confidence intervals for each pattern.

Results

Figure 4 shows the minimal distance determined for each pattern. The mute baseline pattern (S7) yields the smallest average minimal distance. The Kruskal–Wallis test and the Mann–Whitney U test, done at 5% levels of significance, revealed that all the non-baseline sound patterns (S1-S6) gave longer minimal distances than S7. This result validates Hypothesis 1. Based on the statistical tests and the 95 % confidence intervals, we also observed that the sound emission patterns with 5 second timings (S1, S3, and S5) give greater minimal distances than the patterns with 2.5-second timings (S2, S4, and S6). Hypothesis 2 is thus also validated.

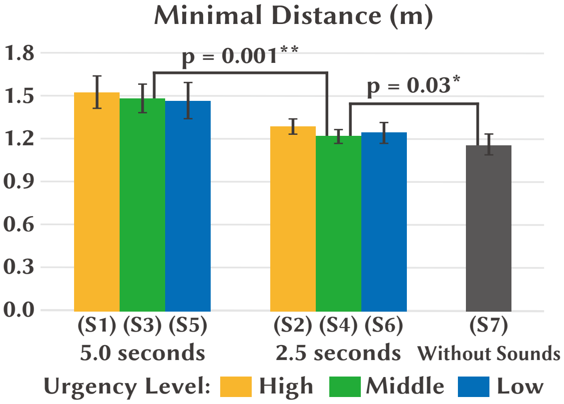

Figure 4: Minimal distances. The bars show the 95 % confidence intervals. *p*: *p*-value of the Mann–Whitney U test done on the minimal distance (∗∗ and ∗ indicate the 0.001 and 0.03 levels of significance, respectively).

We then compared sound patterns with the same emission timing to assess the influence of the urgency levels. We observed no statistically significant difference among patterns with either the 5-second (S1, S3, and S5) or the 2.5-second timings (S2, S4, and S6). We thus rejected Hypothesis 3.

We summarize our findings as follows:

- Sound warnings based on collision prediction influenced pedestrians walking toward away from the suitcase.
- Timings of the sound emissions also affected pedestrian trajectories. Patterns with a 5-second timing deflected pedestrian trajectories more effectively.
- The type of alarm sound appears not to be a significant factor affecting pedestrian trajectories.

Design of the Sound-Emission Policy

Based on the above findings, we designed a sound-emission policy for BBeep (Figure 5), consisting of three stages of sound emissions for preventing collisions. The system emits the following three types of alarm sounds.

(1) A **low-urgency beep** warns of the potential risk of a collision between the blind user and pedestrians within 5 s. This sound was used in our observation study as a *low-urgency sound* S5. We expect this signal to enable pedestrians to divert their path away from the blind user to avoid collision.
(2) An **intermediate-urgency beep** indicates a potential risk of collision within 2.5 s. This sound was used in our

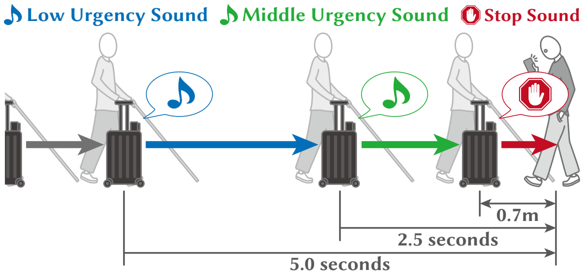

Figure 5: Policy of sound emission.

ID	Gender	Age	Eyesight	Navigation Aid
P1	F	70	Blind	Cane
P2	F	70	Blind	Cane
P3	M	65	Blind	Cane
P4	M	46	Blind	Dog (primary) and Cane
P5	M	42	Blind	Dog (primary) and Cane
P6	M	58	Blind	Cane

Table 2: Demographic information of our participants

observational study as an *intermediate-urgency sound* S4. We also expect this signal to help avoid collision.

(3) A **stop sound** indicates an imminent risk of collision with any obstacle (pedestrian, chair, wall, etc.) located within 70 cm. We expect this signal to prompt the blind user to come to a halt immediately.

We chose the intermediate- and low-urgency sounds for our policy. The higher the urgency level of the signal, the greater the annoyance rating of the sound alert. However, we observed that the urgency level of the sound did not affect the trajectory of oncoming pedestrians. We, therefore, selected two sound alerts with lower urgency and annoyance levels. By using two types of beep sounds, the blind user can know whether or not a pedestrian continues to approach. In addition, to inform the blind user of an obstacle ahead, we use a bell sound that is completely different from the beep sounds. This bell sound is emitted whenever the system detects obstacles located within 70 cm. In our user evaluation, we recommend that the blind user stop advancing immediately upon hearing the bell sound.

6 REAL-WORLD USER EVALUATION

Our main goal was to understand the effectiveness of BBeep in clearing the path for blind travellers in crowded spaces. For that reason, we performed a real-world study where 6 blind participants (Table 2) navigated crowded areas at the Pittsburgh International Airport. In this study, we compared BBeep against two baseline conditions: one notifies only the blind user about collision risks, while the other does not provide any notifications.

Conditions

We equipped our assistive suitcase system with the capabilities to track pedestrians and predict future collisions. Based on this system, we prepared three different interfaces:

Speaker interface (BBeep): This interface represents our proposed system, which emits three types of sounds (low urgency, middle urgency, and stop sounds) for the blind user and other pedestrians through a speaker that is mounted on top of the suitcase.

Headset interface: This interface has the same behavior as BBeep, but instead of using a speaker, provides sounds only to the blind user using bone-conducting headphones (to avoid blocking environmental sound).

No sound interface: The user also carries the suitcase, but this interface never emits sound, representing a blind user navigating by himself without additional aids.

Tasks

We selected several crowded gates where passengers were waiting for boarding in line or in groups. Participants were asked to walk straight along the corridor and go through the crowds until reaching a particular location, where the experimenter would ask them to stop (each task had roughly 20 meters). Participants held the suitcase handle with one hand, and used their white-cane on the other hand (Figure 6). Their goal was to go through the crowds effectively and avoid collisions with other pedestrians. This task enabled us to replicate a very similar setting among different participants and trials, thus enabling a fair comparison among conditions.

Procedure

After obtaining (IRB approved) informed consent from participants, researchers provided an overview of the study and described the three interfaces. A short training session (10 - 15 minutes) was then given to participants until they were familiar with the system alarm sounds and interfaces. Although the volume rate of the speaker interface was fixed for all participants, they were able to adjust the volume in the headset interface to make sure it was comfortable, but audible. During the training session, we explained how to hold the suitcase as it affects the accuracy of collision prediction.

Then, participants were asked to walk five similar routes using three types of interfaces (the speaker and headset interfaces twice and the no sound interface once) in a counterbalanced order. Participants were informed that a researcher would be walking behind them to guarantee their safety as well as other pedestrians' safety (Figure 6). They were also instructed to stop when listening the higher urgency (stop) sound to avoid colliding with pedestrians. The researcher

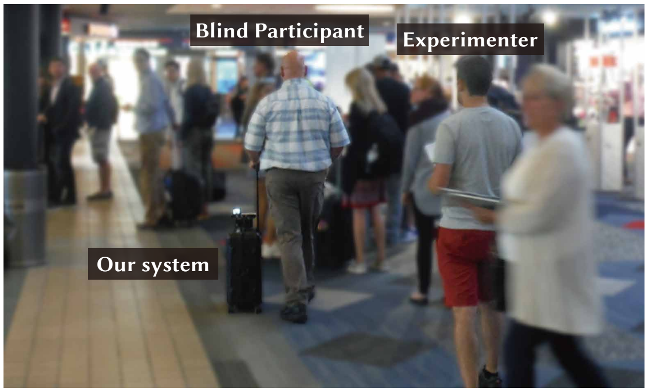

Figure 6: User study at an international airport. Participants walked through crowds holding the handle of an assistive suitcase-shaped system.

did not intervene unless: there was an imminent risk or a deviation from the path. For example, in the latter, the researcher would tell them to slightly adjust their orientation. Also, in case the path was blocked or the participant was confused, the researcher would intervene to help the participant passing that immediate obstacle.

To observe the response of pedestrians to the system, we mounted a GoPro camera on the top of the suitcase.

Metrics

Imminent Collision Frequency and Collision Risk Frequency. To measure how many pedestrians had an imminent risk of collision with the blind user, we defined the number of pedestrians within 70 cm as the *Imminent Collisions Frequency*. In addition, we measured the *Collision Risk Frequency* that indicates how many pedestrians had a risk of collision with the blind user within 5 s. In each task, we counted imminent collisions and risk of collision based on pedestrian detection results and our collision prediction results, respectively. We compared the three conditions quantitatively based on a 95% confidence interval (Table 3). In addition, we compared the two sound conditions (speaker and headset interfaces) using a Wilcoxon signed-rank test with 1% levels of significance.

Risk Continuity Ratio. This metric represents the ratio of pedestrians who had potential risks of collision, and persisted in the users' path until reaching an imminent risk of collision. To calculate the metric, we divided the Imminent Collision Frequency by the Collision Risk Frequency. Smaller values indicate that the system reduces the risk of collision between the blind user and pedestrians. We performed the same analysis described for the previous metrics.

Post-Interview. After completing the tasks, we asked participants to rate a set of sentences using 7-point Likert Items (ranging from 1: strongly disagree, 4: neutral, to 7: strongly agree). The sentences and a summary of the answers are shown in Table 4. Finally, we asked open-ended questions about the advantages and challenges of each interface (speaker and headset). We also asked for suggestions to improve each interface, and in what scenarios would the participants use the Speaker Interface (BBeep).

Results

Quantitative Evaluation. Table 3 reports the imminent collision frequency, collision risk frequency and the risk continuity ratio. We found no significant differences between interfaces concerning the collision risk frequency ($p = 0.8$). On the other hand, our analysis revealed that the speaker interface resulted in significantly ($p = 0.005$) less pedestrians with an imminent risk of collision with the user, when compared to the headset interface. Moreover, a significant difference ($p = 0.009$) in risk continuity ratio shows that the speaker interface was more effective to reduce the number of pedestrians that had a risk of collision with the user.

Table 4 reports the results of six Likert scale questions. Four participants agreed that people cleared the path when they were using the speaker interface rather than the headset interface (Q1). On the other hand, in the other questions, we obtained similar results between two interfaces.

Video Observations. Video recordings enabled us to analyze the behavior of both the blind user and sighted pedestrians, in order to complement our quantitative metrics. We observed that participants would clear the path most of the times after noticing the user. However, participants using the Headset or No Sound interfaces often collided (or had an imminent risk of collision) with pedestrians who were unaware of their presence. In most occasions, pedestrians were either talking in groups or standing in line waiting for boarding. On the other hand, when using the Speaker interface, even in the aforementioned scenarios pedestrians would hear the sound and immediately clear the path for the blind user.

There were, however, five exceptions where pedestrians approached within 70 cm radius of the participant, representing an imminent risk of collision. The reasons for them were: (1) a pedestrian was using headphones and did not hear the sound notification; (2) a blind user changed his walking direction suddenly; and (3) pedestrians tried to clear the path for the blind person, but did not have any anywhere to go (*e.g.,* by being against the wall).

Qualitative feedback. Participants were generally aware that other pedestrians cleared the path when using BBeep, as illustrated by their comments: A1: *"The advantage of the speaker is [that] they [other pedestrians] cleared the path"* (P5); A2: *"People were noticing that I was approaching and people were moving away... giving me the path"* (P4); and A3: *"The biggest advantage is that other people heard it [sound alert] and*

	Collision Risk Frequency			Imminent Collision Frequency			Risk Continuity Ratio		
Interface	Mean and SD	Lower	Upper	Mean and SD	Lower	Upper	Mean and SD	Lower	Upper
Speaker	6.67 ± 3.75	4.55	8.79	**0.41 ± 0.76**	**0.00**	**0.85**	**0.08 ± 0.19**	**0.00**	**0.19**
Headset	5.91 ± 2.25	4.64	7.19	2.00 ± 1.35	1.23	2.76	0.37 ± 0.25	0.22	0.51
No sound	6.67 ± 2.05	5.03	8.30	3.00 ± 1.85	1.52	4.48	0.45 ± 0.21	0.28	0.63

Table 3: Quantitative evaluation of each metric. It presents means and standard deviations (SD), and the lower and upper bounds of 95 % confidence intervals.

No.	Question	P1	P2	P3	P4	P5	P6	mean	SD	Median
Q1	People cleared the path when I was **using the speaker interface**.	4	6	7	7	5	4	5.5	1.26	5.5
	People cleared the path when I was **not using the speaker interface.**	4	3	2	3	4	4	3.33	0.75	3.5
Q2	**The speaker interface** helped me walk comfortably in airports.	7	4	5	6	6	5	5.5	0.96	5.5
	The headset interface helped me walk comfortably in airports.	7	4	3	6	6	6	5.33	1.37	6
Q3	**The speaker interface** is also useful in less crowded places.	5	2	6	6	6	4	4.83	1.46	5.5
	The headset interface is also useful in less crowded places.	5	2	4	5	6	4	4.33	1.25	4.5

Table 4: Likert Items (1: strongly disagree to 7: strongly agree) and a summary of answers.

they would move to get out of the way” (P3). When using the headset interface, participants felt that being quieter was its main advantage: A4: *“It [the headset interface] is more private”* (P1). However, they also had the perception that pedestrians did not clear the path in the same way: A5: *“People don’t notice, so I’m required to say something for them to clear the path.. in comparison to the speaker”* (P4); and A6: *“[The main advantage of the headset interface is that] it’s quiet. ... [The main challenge is that] it didn’t get anybody’s attention”* (P2).

Some participants commented that the usefulness of the speaker interface might depend on the environment. Places where they would use it generally include crowded public spaces and open areas: A7: *“It’s more useful in more crowded places”* (P1); A8: *“It is useful at a grocery store, a shopping mall, and other open areas’* (P5). Still, it was also reported to be useful in less crowded places: A9: *“The speaker interface is also useful in less crowded places, because it doesn’t beep when there are no people. So, I can still take advantage when there is someone with a risk of collision”* (P3). In contrast, using it in quieter environments was found to be inappropriate or to draw too much attention: A10: *“In the airport type of settings, I would probably use the speaker settings, but if I’m in a quiet area where people are expected to be quiet, ... maybe I will not use it”* (P4); and A11: *“I don’t agree to use the speaker interface at places supposed to be quiet like hospitals or libraries, but, in any public environment like airports, train stations, or whatever, the speaker is always gonna be appropriate.”* (P2).

When asked for suggestions, two users said that BBeep should not only beep but also provide information about the surroundings: A12: *“I’m more likely to use the speaker ... but I still want to hear what’s going on through the headset”* (P3); A13: *“[In both interfaces] I want to understand what’s happening. People are in front of me walking, coming, or standing. ... [I recommend] different output. Speaker will notify sighted people. Headset will explain what’s going on.”*(P2).

7 DISCUSSION

Effectiveness of BBeep for Collision Avoidance

The real-world user study showed that BBeep was an effective tool for blind users to prevent collisions with pedestrians. While the number of pedestrians with a low risk of collision (within the next 5 seconds) was very similar among conditions, the number of pedestrians with an imminent risk of collision with the user was significantly lower for BBeep than the Headset condition. These two conditions provide the exact same sound notifications, but use different output sources (*i.e.*, speaker or headset). This result indicates that emitting sound both to the blind user and to nearby pedestrians was effective for clearing the path for the user, and that it was more effective than notifying the user alone. Video observation of the navigation tasks at the airport corroborate these results. Participants traversed crowded areas near the gates and frequently encountered pedestrians who were unaware of them. When walking with BBeep, sighted pedestrians gained immediate awareness of the user’s presence and cleared the path and, in some cases, even prompted their peers to move. Although not always aware of sighted users’ behavior, participants had the perception that BBeep was more effective than the alternatives, as shown by their ratings and comments (A1, A2, and A3).

While the Headset condition was not as effective as BBeep, users took advantage of their knowledge about the collision risk with pedestrians. For instance, P4 started saying *"Excuse me!"* after noticing that the collision risk persisted, while other participants became more effective orienting the suitcase in order to find a path without risks.

Prospective Scenarios for BBeep

We carefully designed our sound emission policy, keeping in mind that social acceptance was crucial for such an approach. In addition, conducting the experiment at the airport enabled participants to understand how would it be to use BBeep in the real-world. Participants' feedback indicates that it is acceptable to use BBeep in crowded, public spaces such as airports, train stations or shopping malls (A7, A8, A10, and A11). Indeed, users' reported comfort in using the suitcase-shaped system (Q2) showed very similar results between the Headset interface and BBeep. In contrast, participants' feedback regarding the use of both interfaces in less crowded places is not consensual (Q3). Still, some participants see advantages in using them since they do not provide notifications unless there are risks of collision (A9). While crowded areas seem appropriate to use BBeep, participants commented that they would not use it in very quiet places where they would attract too much attention or in places where they are supposed to be quiet, such as hospitals or libraries (A10 and A11).

Limitations and Future Work

Reducing the Number of Sound Emissions. The main advantage of the Headset interface was it discreetness, as it does not attract so much attention nor disturb other people (A4 and A6). However, being more private significantly impacted performance. This relation between performance and discreetness suggests that it is important to investigate how to further reduce the number of sound emissions while maintaining its ability to clear the path for blind users. For instance, we observed that sometimes sound notifications were provided even when pedestrians had already noticed the blind user, but did not clear the path immediately. Future solutions may consider using face tracking or gaze estimation techniques [51] to assess whether pedestrians are aware of the blind user, thus reducing the number of sound emissions.

BBeep Acceptability by Sighted Pedestrians. To assess the acceptability of BBeep, it is relevant to investigate not only the impressions of blind users, but also those of sighted pedestrians. However, in this case recruiting sighted people beforehand would prevent us from evaluating BBeep's ability to help clearing the path for the blind user. We aim to further explore sighted people's impressions in the future with a different study design.

Beyond Path Clearing. In order to evaluate the impact of our approach, we focused exclusively on collision avoidance and on the ability to clear the path for the blind user. For that purpose, we used straight-line routes and did not include additional navigational challenges that could affect the results. These design decisions allowed us to run a more controlled experiment, despite being done in a real-world scenario. However, independently traversing complex environments like airports has additional significant challenges such as following a particular route, or gaining knowledge about surrounding Points of Interest (POIs).

The need to convey more informative feedback to the blind user was also mentioned by participants, who wanted to know more details about their surroundings (A12 and A13). One possible extension is to encode distance (or urgency) information continuously instead of using three predetermined levels. A different possibility is to provide the user with additional information that is useful for orientation and mobility. In particular, P2 and P3 suggested to combine the speaker and the headset so that they provide different feedback to the user. They suggested to use BBeep as is, but to describe the environment using the bone-conductive headset. Future directions may include investigating how to combine BBeep with solutions that provide turn-by-turn navigation assistance and/or convey information about relevant POIs in the vicinity of the user [1, 14, 42].

8 CONCLUSION

We proposed an assistive suitcase system, BBeep, that aims to clear the path for blind users when walking through crowded spaces, by notifying both the user and sighted pedestrians about the risks of collision. It provides sound notifications only when needed, based on pedestrian tracking and by predicting their future position in real-time. We first investigated how to convey the sound feedback effectively to sighted pedestrians and designed the sonic notification interface of BBeep. Then, we conducted a real-world user study with visually impaired people in an airport. Results showed that BBeep reduces the number of situations of imminent collision risk when compared to notifying the blind user alone. Moreover, users found BBeep acceptable and appropriate to use in crowded, public spaces such as airports, train stations or shopping malls. Yet, they were more hesitant about using it in places they are supposed to be quiet. In the future, we plan to extend our collision prediction method, by using vision-based attention analysis to reduce the number of unnecessary sound emissions when the pedestrians have already noticed the presence of the blind user.

ACKNOWLEDGMENTS

We thank the Allegheny County Airport Authority and all study participants. This work was sponsored in part by JST CREST (JPMJCR14E1), JST AIP-PRISM (JPMJCR18ZG), NSF NRI award (1637927), NIDILRR (90DPGE0003), Uptake (CMU ML for Social Good fund) and Shimizu Corporation.

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 *Proc. ACM Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, 90–99. https://doi.org/10.1145/2935334.2935361

[2] Tomohiro Amemiya, Jun Yamashita, Koichi Hirota, and Michitaka Hirose. 2004. Virtual leading blocks for the deaf-blind: A real-time way-finder by verbal-nonverbal hybrid interface and high-density RFID tag space. In *Proc. IEEE Conference on Virtual Reality (VR '04)*. IEEE, 165–287. https://doi.org/10.1109/VR.2004.1310070

[3] Mauro Avila Soto, Markus Funk, Matthias Hoppe, Robin Boldt, Katrin Wolf, and Niels Henze. 2017. DroneNavigator: Using Leashed and Free-Floating Quadcopters to Navigate Visually Impaired Travelers. In *Proc. ACM SIGACCESS Conference on Computers and accessibility (ASSETS '17)*. ACM, 300–304. https://doi.org/10.1145/3132525.3132556

[4] JM Bengamin. 1973. The new C-5 laser cane for the blind. In *Conference on the electronic prosthetics, 1973*.

[5] BB Blasch, SJ LaGrow, and WR De l'Aune. 1996. Three aspects of coverage provided by the long cane: Object, surface, and foot-placement preview. *Journal of Visual Impairment and Blindness* 90 (1996), 295–301.

[6] GP Boucek Jr, James E Veitengruber, and Wayne D Smith. 1977. *Aircraft alerting systems criteria study. Volume II. Human factors guidelines for aircraft alerting systems*. Technical Report. BOEING COMMERCIAL AIRPLANE CO SEATTLE WA.

[7] Michael Brock and Per Ola Kristensson. 2013. Supporting blind navigation using depth sensing and sonification. In *Proc. ACM International Joint Conference on Pervasive and ubiquitous computing adjunct publication (UbiComp '13)*. ACM, 255–258. https://doi.org/10.1145/2494091.2494173

[8] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE, 1302–1310. https://doi.org/10.1109/CVPR.2017.143

[9] Ping-Ling Chen, Wafaa Saleh, and Chih-Wei Pai. 2017. Texting and walking: a controlled field study of crossing behaviours and inattentional blindness in Taiwan. *Behaviour & Information Technology* 36, 4 (2017), 435–445. https://doi.org/10.1080/0144929X.2016.1240234

[10] Sakmongkon Chumkamon, Peranitti Tuvaphanthaphiphat, and Phongsak Keeratiwintakorn. 2008. A blind navigation system using RFID for indoor environments. In *Proc. IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON '08)*, Vol. 2. IEEE, 765–768. https://doi.org/10.1109/ECTICON.2008.4600543

[11] Karen Duarte, José Cecílio, and Pedro Furtado. 2014. Easily guiding of blind: Providing information and navigation-smartnav. In *Wireless Internet*. Springer, 129–134. https://doi.org/10.1007/978-3-319-18802-7_18

[12] Judy Edworthy, Sarah Loxley, and Ian Dennis. 1991. Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human factors* 33, 2 (1991), 205–231. https://doi.org/10.1177/001872089103300206

[13] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2012. The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks. In *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, 425–432. https://doi.org/10.1145/2207676.2207735

[14] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2013. Indoor human navigation systems: A survey. *Interacting with Computers* 25, 1 (2013), 21–33. https://doi.org/10.1093/iwc/iws010

[15] José Faria, Sérgio Lopes, Hugo Fernandes, Paulo Martins, and João Barroso. 2010. Electronic white cane for blind people navigation assistance. In *Proc. IEEE Conference on World Automation Congress (WAC '10)*. IEEE, 1–7.

[16] Vítor Filipe, Filipe Fernandes, Hugo Fernandes, António Sousa, Hugo Paredes, and João Barroso. 2012. Blind navigation support system based on Microsoft Kinect. *Procedia Computer Science* 14 (2012), 94–101. https://doi.org/10.1016/j.procs.2012.10.011

[17] Masahiro Furukawa, Hiromi Yoshikawa, Taku Hachisu, Shogo Fukushima, and Hiroyuki Kajimoto. 2011. "Vection field" for pedestrian traffic control. In *Proc. ACM International Conference on Augmented Human (AH '11)*. ACM, 19. https://doi.org/10.1145/1959826.1959845

[18] Thomas Gallagher, Elyse Wise, Binghao Li, Andrew G Dempster, Chris Rizos, and Euan Ramsey-Stewart. 2012. Indoor positioning system based on sensor fusion for the blind and visually impaired. In *Proc. IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN '12)*. IEEE, 1–9. https://doi.org/10.1109/IPIN.2012.6418882

[19] Christian Gonzalez, Bridget A Lewis, Daniel M Roberts, Stephanie M Pratt, and Carryl L Baldwin. 2012. Perceived urgency and annoyance of auditory alerts in a driving context. 56, 1 (2012), 1684–1687. https://doi.org/10.1177/1071181312561337

[20] Ellen C Haas and Judy Edworthy. 1996. Designing urgency into auditory warnings using pitch, speed and loudness. *Computing & Control Engineering Journal* 7, 4 (1996), 193–198. https://doi.org/10.1049/cce:19960407

[21] Juan David Hincapié-Ramos and Pourang Irani. 2013. CrashAlert: enhancing peripheral alertness for eyes-busy mobile interaction while walking. In *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, 3385–3388. https://doi.org/10.1145/2470654.2466463

[22] Hsieh-Chang Huang, Ching-Tang Hsieh, and Cheng-Hsiang Yeh. 2015. An indoor obstacle detection system using depth information and region growth. *Sensors* 15, 10 (2015), 27116–27141. https://doi.org/10.3390/s151027116

[23] Andreas Hub, Joachim Diepstraten, and Thomas Ertl. 2004. Design and development of an indoor navigation and object identification system for the blind. In *Proc. ACM SIGACCESS Conference on Computers and accessibility (ASSETS '04)*. ACM, 147–152. https://doi.org/10.1145/1029014.1028657

[24] Kiyohide Ito, Makoto Okamoto, Junichi Akita, Tetsuo Ono, Ikuko Gyobu, Tomohito Takagi, Takahiro Hoshi, and Yu Mishima. 2005. CyARM: an alternative aid device for blind persons. In *Proc. ACM CHI Extended Abstracts on Human Factors in Computing Systems (CHI '05)*. ACM, 1483–1488. https://doi.org/10.1145/1056808.1056947

[25] Beomjoon Kim and Joelle Pineau. 2016. Socially Adaptive Path Planning in Human Environments Using Inverse Reinforcement Learning. *International Journal of Social Robotics* 8, 1 (2016), 51–66. https://doi.org/10.1007/s12369-015-0310-2

[26] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. 2016. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In *Proc. ACM Annual Symposium on Applied Computing (AC '16)*. ACM, 604–611. https://doi.org/10.1145/2851613.2851716

[27] Kris M. Kitani, Brian D. Ziebart, James Andrew Bagnell, and Martial Hebert. 2012. Activity Forecasting. In *Proc. IEEE European Conference on Computer Vision (ECCV '12)*, Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid (Eds.). Springer Berlin Heidelberg, 201–214. https://doi.org/10.1016/B978-0-12-809276-7.00014-X

[28] Vladimir Kulyukin, Chaitanya Gharpure, John Nicholson, and Grayson Osborne. 2006. Robot-assisted wayfinding for the visually impaired

in structured indoor environments. *Autonomous Robots* 21, 1 (2006), 29–41. https://doi.org/10.1007/s10514-006-7223-8

[29] Bing Li, J Pablo Munoz, Xuejian Rong, Jizhong Xiao, Yingli Tian, and Aries Arditi. 2016. ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind. In *Proc. IEEE European Conference on Computer Vision (ECCV '16)*. Springer, 448–462. https://doi.org/10.1007/978-3-319-48881-3_31

[30] Shachar Maidenbaum, Shlomi Hanassy, Sami Abboud, Galit Buchs, Daniel-Robert Chebat, Shelly Levy-Tzedek, and Amir Amedi. 2014. The "EyeCane", a new electronic travel aid for the blind: Technology, behavior & swift learning. *Restorative neurology and neuroscience* 32, 6 (2014), 813–824.

[31] Roberto Manduchi, Sri Kurniawan, and Homayoun Bagherinia. 2010. Blind guidance using mobile computer vision: A usability study. In *Proc. ACM SIGACCESS Conference on Computers and accessibility (ASSETS '10)*. ACM, 241–242. https://doi.org/10.1145/1878803.1878851

[32] Dawn C Marshall, John D Lee, and P Albert Austria. 2007. Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human factors* 49, 1 (2007), 145–157. https://doi.org/10.1518/001872007779598145

[33] Edward Marshall and Sue Baker. 1995. Alarms in nuclear power plant control rooms: current approaches and future design. In *Human factors in alarm design*. Taylor & Francis, Inc., 183–191.

[34] Christina Meredith and Judy Edworthy. 1995. Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of advanced nursing* 21, 1 (1995), 15–20. https://doi.org/10.1177/001872089303500408

[35] Madoka Nakajima and Shinichiro Haruyama. 2012. Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing. In *Proc. IEEE International Conference on Communications in China (ICCC '12)*. IEEE, 524–529. https://doi.org/10.1109/ICCChina.2012.6356940

[36] Helen Petrie, Valerie Johnson, Thomas Strothotte, Andreas Raab, Steffi Fritz, and Rainer Michel. 1996. MoBIC: Designing a travel aid for blind and elderly people. *The Journal of Navigation* 49, 1 (1996), 45–52. https://doi.org/10.1145/191028.191051

[37] Huy-Hieu Pham, Thi-Lan Le, and Nicolas Vuillerme. 2016. Real-time obstacle detection system in indoor environment for the visually impaired using microsoft kinect sensor. *Journal of Sensors* 2016 (2016). https://doi.org/10.1155/2016/3754918

[38] Ioannis Politis, Stephen A Brewster, and Frank Pollick. 2014. Evaluating multimodal driver displays under varying situational urgency. In *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 4067–4076. https://doi.org/10.1145/2556288.2556988

[39] Lisa Ran, Sumi Helal, and Steve Moore. 2004. Drishti: an integrated indoor/outdoor blind navigation system and service. In *Proc. the Second IEEE Annual Conference on Pervasive Computing and Communications (PerCom '04)*. IEEE, 23–30. https://doi.org/10.1109/PERCOM.2004.1276842

[40] Joseph Redmon and Ali Farhadi. 2017. YOLO9000: Better, Faster, Stronger. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE, 6517–6525. https://doi.org/10.1109/CVPR.2017.690

[41] Shadan Sadeghian Borojeni, Susanne CJ Boll, Wilko Heuten, Heinrich H Bülthoff, and Lewis Chuang. 2018. Feel the Movement: Real Motion Influences Responses to Take-over Requests in Highly Automated Vehicles. In *Proc. ACM CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 246. https://doi.org/10.1145/3173574.3173820

[42] Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In *Proc. ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. ACM, 270–279. https://doi.org/10.1145/3132525.3132535

[43] Shraga Shoval, Johann Borenstein, and Yoram Koren. 1994. Mobile robot obstacle avoidance in a computerized travel aid for the blind. In *Proc. IEEE International Conference on Robotics and Automation*. IEEE, 2023–2028.

[44] Shraga Shoval, Johann Borenstein, and Yoram Koren. 1998. The Navbelt-A computerized travel aid for the blind based on mobile robotics technology. *IEEE Transactions on Biomedical Engineering* 45, 11 (1998), 1376–1386. https://doi.org/10.1109/10.725334

[45] Yu-Chih Tung and Kang G Shin. 2018. Use of Phone Sensors to Enhance Distracted Pedestrians' Safety. *IEEE Transactions on Mobile Computing* 17, 6 (2018), 1469–1482. https://doi.org/10.1109/TMC.2017.2764909

[46] Iwan Ulrich and Johann Borenstein. 2001. The GuideCane-applying mobile robot technologies to assist the visually impaired. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 31, 2 (2001), 131–136.

[47] William R Wiener, Richard L Welsh, and Bruce B Blasch. 2010. *Foundations of orientation and mobility*. Vol. 1. American Foundation for the Blind.

[48] Takuma Yagi, Karttikeya Mangalam, Ryo Yonetani, and Yoichi Sato. 2018. Future Person Localization in First-Person Videos. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '18)*.

[49] Cang Ye, Soonhac Hong, Xiangfei Qian, and Wei Wu. 2016. Co-robotic cane: A new robotic navigation aid for the visually impaired. *IEEE Systems, Man, and Cybernetics Magazine* 2, 2 (2016), 33–42. https://doi.org/0.1109/MSMC.2015.2501167

[50] Limin Zeng, Markus Simros, and Gerhard Weber. 2017. Camera-based mobile electronic travel aids support for cognitive mapping of unknown spaces. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, 8. https://doi.org/10.1145/3098279.3098563

[51] Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. 2017. Mpiigaze: Real-world dataset and deep appearance-based gaze estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017). https://doi.org/10.1109/TPAMI.2017.2778103