Achieving Practical and Accurate Indoor Navigation for People with Visual Impairments

Dragan Ahmetovic
Carnegie Mellon University
dragan1@cmu.edu

Masayuki Murata
IBM Research-Tokyo
muratams@jp.ibm.com

Cole Gleason
Carnegie Mellon University
cgleason@cs.cmu.edu

Erin Brady
Indiana University - Purdue University Indianapolis
brady@iupui.edu

Hironobu Takagi
IBM Research-Tokyo
takagih@jp.ibm.com

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

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

ABSTRACT

Methods that provide accurate navigation assistance to people with visual impairments often rely on instrumenting the environment with specialized hardware infrastructure. In particular, approaches that use sensor networks of Bluetooth Low Energy (BLE) beacons have been shown to achieve precise localization and accurate guidance while the structural modifications to the environment are kept at minimum. To install navigation infrastructure, however, a number of complex and time-critical activities must be performed. The BLE beacons need to be positioned correctly and samples of Bluetooth signal need to be collected across the whole environment. These tasks are performed by trained personnel and entail costs proportional to the size of the environment that needs to be instrumented.

To reduce the instrumentation costs while maintaining a high accuracy, we improve over a traditional regression-based localization approach by introducing a novel, graph-based localization method using Pedestrian Dead Reckoning (PDR) and particle filter. We then study how the number and density of beacons and Bluetooth samples impact the balance between localization accuracy and set-up cost of the navigation environment. Studies with users show the impact that the increased accuracy has on the usability of our navigation application for the visually impaired.

CCS Concepts

•Social and professional topics → **People with disabilities; •Human-centered computing** → **Accessibility technologies; User studies; •Computer systems organization** → *Sensor networks;* **•Information systems** → *Location based services;*

Keywords

Visual Impairments, Navigation Assistance, Indoor Localization

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

W4A 2017, April 02 - 04, 2017, Perth, Western Australia, Australia

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ISBN 978-1-4503-4900-0/17/04…$15.00

DOI: http://dx.doi.org/10.1145/3058555.3058560

1. INTRODUCTION

To provide navigation assistance to individuals with visual impairments, it is imperative to accurately localize them within the environment. While traditional means of localization use GPS signal to provide rough estimates of a user's position outdoors [Ahmetovic et al. 2015, Kacorri et al. 2016], it is challenging to obtain the level of accuracy needed to guide a blind person in indoor environments.

Existing technologies such as networks of laser range finders [Fod et al. 2002] or cameras [Dockstader and Tekalp 2001] can localize people within an environment. However, such systems require large investments in terms of hardware, installation and calibration costs, which makes them impractical for many building owners who may wish to make their building accessible to the blind. Scalable sensor networks have been proposed to provide navigation assistance for individuals with visual impairments at low cost and without restructuring the environment [Gonzalez et al. 2007, Saito et al. 2007]. In particular, NavCog is an open source software that relies on Bluetooth low-energy (BLE) beacons installed in an environment and commodity smartphones to provide accurate navigation assistance to individuals with visual impairments [Ahmetovic et al. 2016b].

To instrument an environment with NavCog infrastructure, it is necessary to correctly position the BLE beacons in the environment (See Figure 1) and to collect samples ("fingerprints") of Bluetooth signals from the installed beacons across the whole environment.

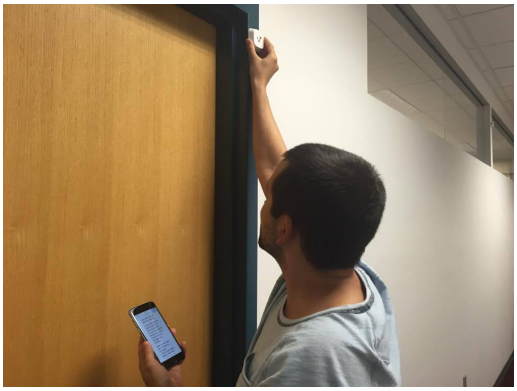
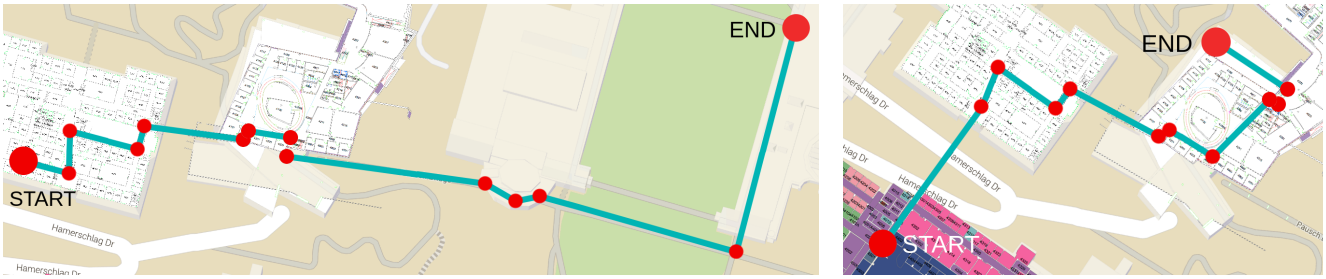


Figure 1: Beacon installation is performed by expert personnel.



(a) Route 1: 390m, 14 nodes, 12 edges

(b) Route 2 - 230m, 12 nodes, 10 edges

Figure 2: Layout of the two paths used during the user experiments. Blue lines represent the straight edges that define the paths, while the red dots signal the turning points that connect consecutive edges. The bigger red dots define the starting and ending nodes.

These tasks require prior expertise with the system and therefore they are performed by trained personnel [Ahmetovic et al. 2016b]. Clearly, the cost related to the installation of the NavCog system increases with the size of the instrumented environment: installing the system in large environments can become problematic as it requires a sufficient number of experts to perform the installation and a sufficient number of Bluetooth beacons to cover the whole area. Thus, it is important to contain the cost related to the installation of the system while providing adequate accuracy during localization and navigation assistance of individuals with visual impairments.

We improve the localization method described in [Ahmetovic et al. 2016b] through the use of pedestrian dead reckoning (PDR) based on the data from the Inertial Motion Unit (IMU) on the user's smartphone, and the information on the position of the BLE beacons. These two information, integrated using a Particle Filtering framework, allow us to achieve higher levels of localization accuracy then what was previously possible with the NavCog system.

To facilitate the introduction of this technology to new environments, we perform extensive evaluations of the effect of sensor placement on localization accuracy and installation costs (time and effort). In particular, we investigate how to achieve different localization accuracy levels with the proposed technique by varying beacon density, fingerprinting resolution, and the number of signal fingerprints gathered at each position. We show that our proposed system achieves a localization accuracy of 0.68m using only a small number of beacons (to a minimum of 1 beacon every 6m).

To show the effectiveness of our system for the real-world navigation tasks for the blind, we evaluate the impact of increased localization accuracy to assist in the navigation task. We asked 6 blind tests subjects to navigate through an unfamiliar environment using NavCog across two routes, 200m and 400m long, shown in Figure 2. Our test show that the quality of the navigation experience was greatly enhanced by the increase in localization accuracy.

2. RELATED WORK

Assistive technologies for helping people with visual impairments to navigate in unexplored environments can be divided in two categories: tools that augment the user's sensing capabilities, and assistive technologies integrated in the environment. The first category aims to improve sensing capabilities of blind pedestrians with wearable or carried tools that can be used to receive additional information from the environment. The most common aid in this group is the long cane [Blasch et al. 1996], used by people with visual impairments to expand their haptic exploration area, and to notify sighted pedestrians of the user's visual impairment. Other tools involve augmentations or replacements of the long cane, such as laser canes [Benjamin and Malvern 1973], ultrasonic sensing devices [Pressey 1977] or depth sensors [Brock and Kristensson 2013]. Mobile phone solutions that perform visual recognition of environment features have also been proposed [Manduchi et al. 2013, Ahmetovic et al. 2014, Fiannaca et al. 2014, Peraković et al. 2015].

The second group augments the environment by adding auditory or haptic cues that a blind pedestrian can perceive and use for improving the understanding of the environment. One of the first assistive technologies of this kind is the tactile paving [Iwahashi 1983], which consists in integrating paving tiles with haptic patterns that people with visual impairments can sense and follow with feet or white cane. Braille tags are another assistive tool in this group, commonly used to inform the user in proximity of doors or on the elevators [Kleege 2006]. Audible street signals, that notify blind pedestrian when they can traverse a crossing, or acoustic floor alerts on modern elevators also belong to this category [Poulsen 1982, Legge et al. 2013].

Assistive technologies combining the two approaches (sensors integrated in the environment and mobile sensing tools carried by the user) can be used to localize the user without the need of direct interaction. The most common solution in this group is GPS positioning, which provides a localization accuracy within tens of meters and it is limited to outdoor use [Manduchi and Kurniawan 2012, Ahmetovic et al. 2015, Ahmetovic et al. 2016, Kacorri et al. 2016]. A solution that uses sporadic direct sensing is Navatar [Fallah et al. 2012]. It leverages the smartphones' inertial motion unit (IMU) sensors for estimating the user's position, which is periodically confirmed by the user by identifying surrounding haptic landmarks. RFID tags [Faria et al. 2010, Chumkamon et al. 2008, Amemiya et al. 2004] installed in the pavement can be detected with a white cane augmented with an RFID reader while the user is moving, thus providing real time localization. Visible Light Communication (VLC) technology [Nakajima and Haruyama 2012] can be used in a similar fashion: VLC-augmented LED lights installed in the environment can be perceived by the smartphone's camera and therefore provide localization information.

Solutions that do not require the use of proprietary hardware or extensive modifications to the environment entail a lower installation and distribution cost. In particular, we are interested in assistive technologies that use off-the-shelf and widespread devices that a blind person might already have, such as smartphones [Gonzalez et al. 2007, Saito et al. 2007]. Wi-Fi signals already existing in the environment have been used to estimate a smartphone's position with an accuracy of about 2m [Hilsenbeck et al. 2014]. Since the Wi-Fi coverage in the environment is not tuned for localization, these approaches do not provide predictable and consistent localization accuracy. Battery powered bluetooth low-energy (BLE) beacons can also be used to localize a smartphone sensing device in

a similar fashion as with the Wi-Fi signal, but they can be installed almost anywhere, as needed, without requiring a high installation and maintenance cost.

NavCog [Ahmetovic et al. 2016a, Ahmetovic et al. 2016b] uses BLE beacons to provide sub-meter localization and navigation assistance for individuals with visual impairments. The method used in NavCog consists of gathering unique "fingerprints" of Received Signal Strength Indication (RSSI) measurements of multiple BLE beacons in a single point of space. Once the RSSI fingerprints for each different point of the environment have been collected, a sensing device can then be localized by comparing its RSSI readings to the prerecorded fingerprints. To avoid interference on the bluetooth signal caused by passer-by, which can impair the localization accuracy [Ryckaert et al. 2004], the NavCog installation procedure requires for the BLE beacons to be positioned high on the walls so that the signal reaches the user's mobile device without traversing other people. In practice, we notice that the system is robust with respect to the interference caused by passer-by and objects in the environment. A formal evaluation of the impact of human and object-caused interference on the localization accuracy of the system is out of the scope of this contribution, but it will be considered as a future extension of this work.

NavCog also approximates the environment to a graph of one-dimensional line segments, based on the "Manhattan world assumption" [Coughlan and Yuille 2000], since the navigation environment is mostly composed by straight paths. This approximation, shown in Figure 3, serves to limit the number of beacons and signal fingerprints required to achieve a high localization accuracy.

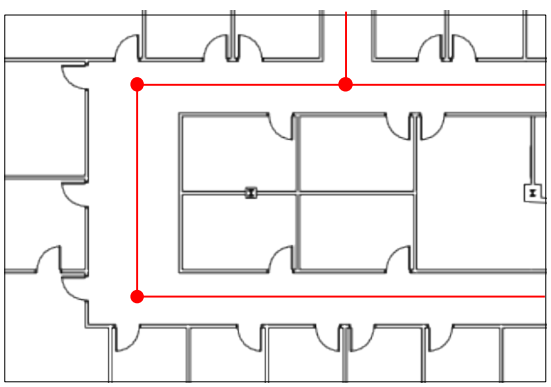


Figure 3: One-dimensional representation of a navigation field.

We further improve the approach described in [Ahmetovic et al. 2016b] to achieve sub-meter accuracy in our 2 different testing environments using as little as 1 beacon every 6m. We achieve this using a multi-modal probabilistic state estimation algorithm and a Particle Filtering framework which maintains a non-parameteric estimate of the user's position. The user's position is updated based on beacon RSSI and an incremental motion estimate based on pedestrian dead reckoning (PDR) using the smartphone IMU data.

Existing work examine how deployment parameters and signal settings influence localization accuracy [Faragher and Harle 2015]. However, to the best of our knowledge, no previous work addresses the issue of how to deploy BLE beacons and collect signal fingerprints to obtain a desired accuracy level while limiting the installation workload. We study the relationship between the number of beacons installed, number of fingerprints, and signal fingerprint resolution with respect to the expected localization accuracy. Finally, we evaluate the localization capabilities of our approach in a real world environment with 6 test subjects with visual impairments.

3. LOCALIZATION METHOD

The localization provided by the NavCog system is constrained to a graph-based representation of the environment that models the navigation field as a graph of one-dimensional line segments (edges). As stated previously, this approximation drastically reduces the workload required to instrument a large environment (navigation field), but it is also compatible with the navigation strategy that individuals with visual impairments habitually use: following walls and other physical cues that indicate the direction of a path. Besides straight paths, some open areas or curved paths with walls can also be approximated as one dimensional edges.

This one dimensional approach, however, is not always suitable to model an environment. For example, it is a poor approximation for areas that cannot be represented as one dimensional paths, like unconstrained open areas without any reference system (e.g., no walls or curbs to follow). In those particular cases, a user navigating without a two-dimensional reference frame may veer away from the intended path [Williams et al. 2014]. In future work, we will investigate how to tailor a navigation field to combine our proposed approach with traditional 2-d localization methods by considering the needs of each portion of the environment.

The one dimensional navigation model achieves accurate localization with a small number of beacons, but, to further improve over the technique proposed in [Ahmetovic et al. 2016b], we integrate two new components to the position estimation method: a probability distribution of RSSI readings from BLE beacons and a PDR motion model based on sensor readings from the smartphone's accelerometer and gyroscope.

3.1 RSSI Probability distribution

A smartphone in the range of a beacon measures its Received Signal Strength Indicator (RSSI), that is the decibel level of the transmitted signal. As the RSSI from a beacon changes with the relative position between the beacon and the smartphone, readings from multiple beacons create a specific "Fingerprint" of the beacon signals for a given position. It is then possible to localize a sensing smartphone device by comparing the RSSI detected from surrounding beacons with the pre-recorded fingerprints. If we know the positions of the beacons in the environment, we can also model the probability distribution of the RSSI of the beacons across the environment from RSSI samples collected after beacon installation. This approach allows us to achieve reasonable prediction of BLE beacon RSSI at any given location with just a few fingerprint samples as training data (See Figure 4), resulting in a more accurate localization. The probability distribution of the RSSI samples is modelled by a Gaussian distribution with a mean $\mu(x_t)$ and a standard deviation $\sigma(x_t)$ based on the mean of the RSSI at location x_t (modelled by the kernel ridge regression [Murphy 2012]).

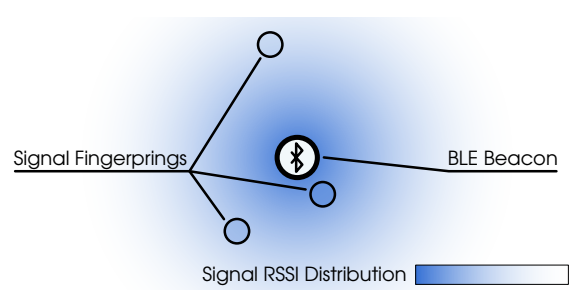


Figure 4: The Probability distribution of the Bluetooth RSSI of a BLE beacon is computed using just a few signal fingerprints.

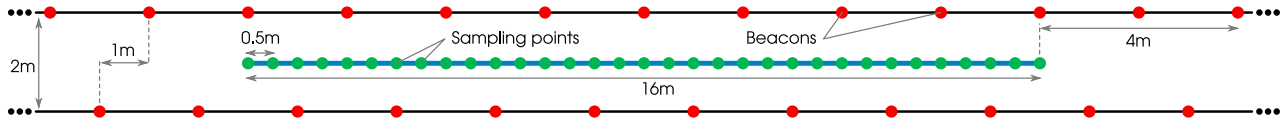


Figure 5: The experimental setting edge used for localization accuracy evaluation.

3.2 Pedestrian Motion Model

The pedestrian motion model predicts the user's location by pedestrian dead reckoning (PDR) using the inertial sensors (accelerometers and gyroscopes) present on modern smartphones. The inertial measurements are used to detect the user's motion state (*i.e.,* moving or stopped) and direction (See Figure 6). A common PDR approach is to count individual steps by applying peak detection to accelerometer measurements [Hilsenbeck et al. 2014]. In our implementation, the user's motion state is detected by computing the standard deviation of the magnitude of accelerometer measurements in a short time window and thresholding the standard deviation by a reference threshold, as proposed in [Brajdic and Harle 2013]. Note that we only used the accelerometer to distinguish if a user is walking or stopped, as developing a robust step detector for different gaits without training the model for each user is difficult. Instead, the user's velocity is estimated based on beacon RSSI localization.

During navigation a user is expected to hold a smartphone in such way that the phone's orientation is approximately parallel to the user's direction. Thus, the phone's orientation (black arrows in Figure 6) can be viewed as the user's walking direction. To represent the user's orientation relative to the traversed edge, the angle between user's walking direction and the edge is computed using gyroscope data. The direction angle is used to understand if the user has correctly modified the direction when turning between adjacent edges, and to estimate the user's velocity.

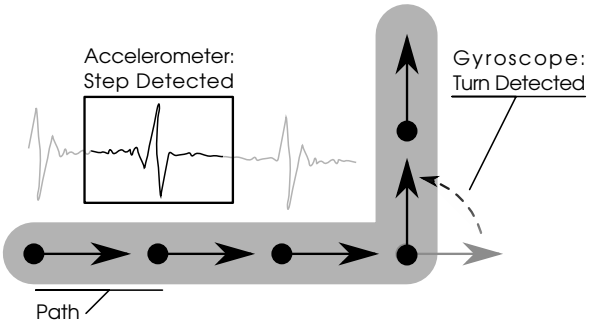


Figure 6: Pedestrian Dead Reckoning: Accelerometers are used to detect steps while the gyroscope tracks the orientation.

4. LARGE SCALE DEPLOYMENT

The localization using the NavCog system, with the improvements described in the previous section, allows us to achieve a high localization accuracy even with few beacons, as it uses a PDR model of the user's movement to estimate the user's position in addition to beacon-based localization. Representing our navigation field as a graph of one-dimensional edges also allows us to reduce the surveying and sampling effort without loss of accuracy, as the sampling points have to be collected only along the edges composing the navigation field. Another advantage of the graph-based representation is that edges are connected only through single points (nodes). This makes the coordinate system of each edge independent from others, which allows us to a) gather sampling data without defining or enforcing a common coordinate system among the entire navigation field and b) perform localization only on the currently traversed edge and switch to the localization on the next edge when the connection node between the edges is reached.

In the following section we study the localization accuracy achieved by our technique in a test environment under different set-up conditions. We show how the density of BLE beacons, sampling resolution, and number of samples per point influence the localization accuracy. We compare the accuracy of the proposed method to the k-nearest neighbor (k-NN) regression[Muja and Lowe 2009] approach used in the previous NavCog version [Ahmetovic et al. 2016b]. We then define the minimum required workload metric that combines the sampling resolution and number of samples per point to express the effort of instrumenting an edge in terms of seconds per meter. We propose a method for defining new navigation fields by balancing the number of beacons needed and the minimum required instrumentation workload to achieve a target accuracy level. As a validation of our method, we consider 3 other edges in two different navigation fields and predict their accuracy based on the number of beacons, sampling resolution, and samples per point used. We show how the predicted accuracy is consistent with the actual accuracy obtained for the test edges.

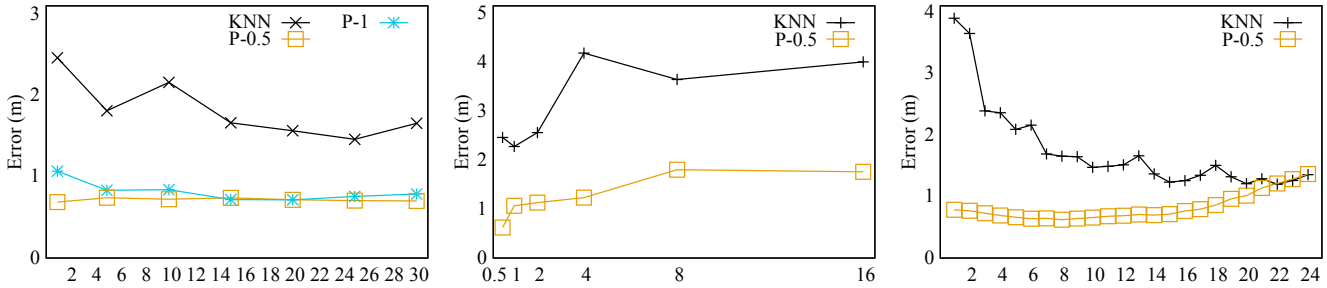
4.1 Experimental Setting

For the localization accuracy evaluation we use 24 Kontakt.io [1] Smart Beacons, tuned for a transmission strength of -12dBm (20 meter range) and a transmission interval of 100ms. For the k-NN regression approach used as a baseline to evaluate localization accuracy evaluation of our proposed solution, we consider a weighted k-NN using euclidean distance and $k = 3$. In our proposed approach, to prevent the variety of particles from diminishing, we empirically tuned the value of the smoothing coefficient α to 0.3.

We set up 4 testing edges. A 16m long edge in a 2 meter wide corridor is used for localization accuracy evaluation (evaluation edge, shown in Figure 5). The edge has been marked every 0.5m and the beacons were positioned at every meter along alternating sides of the corridor, starting from 4 meters before the edge and ending 4 meters after the edge. The other 3 edges (lengths 15m, 16m, and 21m) are used to validate our method for defining set up parameters for navigation fields based on the desired accuracy level (validation edges). These edges were marked every 4m, 16m, and 1m, and beacons were deployed respectively every 4m, 8m, and 4m. On the first edge, 30 RSSI signal samples were captured for each marked sampling point. On the second and third edges, 20 and 10 samples per point were collected, respectively.

We collected 10 egocentric videos of a test subject traversing the evaluation edge and 1 video for each validation edge. The videos were captured with a 1280x720 pixels resolution at a minimum of 120fps. Corresponding beacon, accelerometer and gyroscope data were captured at 1fps, 100fps, and 20fps respectively. From videos we extract the frames and compute the corresponding timestamps in which the test subject traversed marked positions of the edge.

[1] http://kontakt.io/



(a) Samples per sampling point (b) Sampling point resolution (c) Number of beacons

Figure 7: Impact of the number of beacons, samples per point and sampling resolution on the localization accuracy.

4.2 Localization Accuracy Evaluation

To evaluate the accuracy of the localization, we consider the positions of the test subject at marked points while traversing the evaluation edge as extracted from video frames. We compute the position estimation on marked points based on beacon, accelerometer and gyroscope readings at that given moment. As our accuracy metric, we then measure the distance between the estimated position and the ground truth.

By excluding data at run time, we can artificially repeat the computation, varying the number of beacons, sampling resolution, and number of samples per point. We apply the same procedure for the k-NN regression, which only considers beacon data for the position estimation, and compare the accuracy obtained in the two approaches.

Unless noted otherwise, the following results assume that 1 beacon were used every 3m (a total of 8 beacons), with a sampling resolution of 0.5m and 30 samples per point. In Figure 7, the k-NN regression localization is labeled as **KNN**, while the proposed solution is labeled as **P-0.5** when the localization is performed with the default sampling resolution of 0.5m and **P-1** when the sampling resolution parameter is set to 1m.

The number of samples per sampling point does influence the k-NN approach, for which it causes an apparently linear reduction in localization error: from 2.5m with a single sample to 1.5m error with 30 samples, as seen in Figure 7(a). A similar effect can be noticed on our proposed solution, when the resolution is set to 1m (from 1.2m to 0.8m). This result can be observed at resolutions higher than 1m. The variation, however, cannot be noticed at a resolution of 0.5m, which maintains a consistent localization error of 0.68m.

The sampling point resolution (see Figure 7(b)) impacts both techniques significantly. For the k-NN approach, the localization error increases slightly between 0.5m and 2m resolution, from 2.2m to 2.5m error respectively. Above a 2m resolution, the error increases greatly to about 4m. For the proposed approach the increase is noticeable even between 0.5m and 1m resolution. The average error in the first case is about 0.68m while in the second it is already up to 1m. It is worth noticing, however, that after that limit, the error increases gracefully, and even with 16m resolution, the localization error is at most 2m.

For evaluating how the localization accuracy varies based on the amount of beacons installed, we applied a recursive procedure (see Algorithm 1). We evaluate the accuracy on the path with all the available beacons. We then disable one beacon at a time and evaluate the accuracy without that beacon. We repeat the procedure recursively without the beacon having the lowest impact on the localization accuracy until only one beacon remains.

Algorithm 1 Beacons evaluation procedure

Input: Set of beacons B.
Output: a set Z of results $z_i = \{B_i, r_i\}$, where B_i is a set of beacons and r_i is the corresponding accuracy result.
Method:

```
Z ← ∅
function RECURSIVEBEACONS(B)
    r ← evaluateAccuracy(B)
    z ← {B, r}
    Z ← Z + z
    z_best ← NULL
    for all (beacon b ∈ B) do
        B_i ← B − b
        r_i ← evaluateAccuracy(B_i)
        z_i ← {B_i, r_i}
        if r_i > z_best(r) then
            z_best ← z_i
        end if
    end for
    B_best ← z_best(B)
    return RECURSIVEBEACONS(B_best)
end function
return Z
```

Figure 8 shows at what stage the beacons are pruned by the selection procedure. The score is defined as the stage at which the beacon is pruned divided by the number of beacons. The figure suggests that the beacon selection procedure is likely to preserve beacons on the extremities of the edge, which are fundamental for correctly localizing the device at the beginning of the localization. Once the device is correctly localized, PDR can efficiently maintain the localization, therefore the presence of beacons in the middle part of the edge is less important.

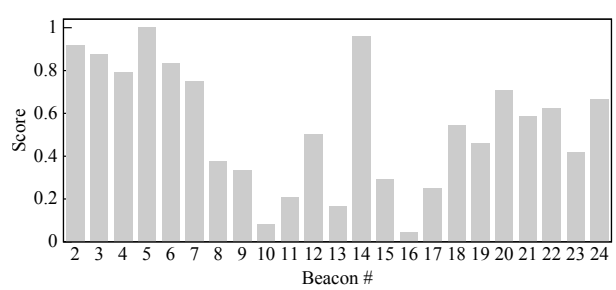
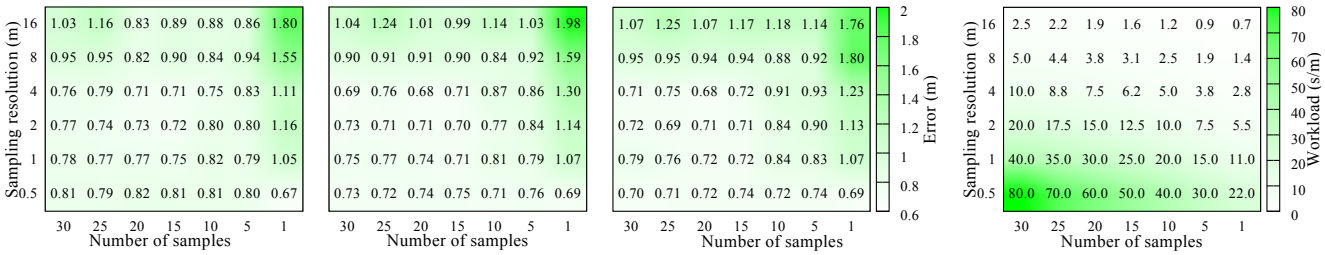


Figure 8: The selection procedure favors beacons on edge ends.



(a) Accuracy - 1 beacon /6m (b) Accuracy - 1 beacon /4m (c) Accuracy - 1 beacon /3m (d) Workload

Figure 9: Average localization error and workload with respect to the sampling resolution and samples per point.

The results in Figure 7(c) show that the proposed approach achieves lower localization error than the k-NN regression, and reaches best results with as little as 6 beacons (one every 4m). Sub-meter accuracy is consistently reached with even one single beacon. Without the PDR motion model, the number of beacons greatly impacts the localization accuracy. In this case, best results are obtained with 16 beacons, that is one beacon every 1.5m. In this best case scenario, the average error is 1.2m for the k-NN regression baseline. The results shown refer to a sampling resolution of 0.5m and 30 samples per point, but similar results are obtained for other resolutions and numbers of samples.

4.3 Minimum Required Workload

When collecting RSSI samples in the environment, the sampling resolution r and the number of samples s, collected for each sampling point, determine how much time it is needed to collect data for a given edge. Based on these two parameters, we define workload w as the time required to sample one meter of an edge. For collecting data, we first need to measure the length of the edge we want to sample. As this procedure needs to be done anyways, independently from the other parameters, we exclude it from the workload computation.

Once we measure the edge, we mark the positions on which to collect data samples. Assuming a constant time $m = 10$s (as an example) for marking each point, and a number of points per meter $n = \frac{1}{r}$, the marking workload is $w_m = \frac{m}{r}$. Once the positions are marked, the sampling workload w_s is the product of the number of sampling points per meter n and samples collected for each point s, that is $w_s = \frac{s}{r}$. The total workload w can therefore be computed as the sum of w_s and w_m, that is $w = \frac{s+m}{r}$.

Figure 9(d) shows how the workload varies with respect to the sampling resolution and number of samples per point, while Figures 9(a), 9(b), and 9(c) show how the relationship between r and s influences the average localization error with one beacon every 6 4 and 3 meters respectively. Workload maps in cases with beacons distanced less than 3m can be omitted as they show no improvements (as already shown in Figure 7(c)). Similar maps have been produced also for k-NN based regression.

As an example, we will consider the situation with 8 beacons in Figure 9(c) (similar considerations can be made in the other two cases). If we wish to achieve a 0.75m accuracy, we have a few choices. We can either take 5 samples at 0.5m resolution, 15 samples at 2m resolution, or 25 samples at 4m resolution. It is also possible to achieve the desired minimum accuracy with different settings, but those will require more workload than the already considered elements along at least one axis. For example, we disregard the option of 20 samples at 1m resolution since it will surely require more workload than 15 samples at 2m resolution. For the remaining options we consider the workload map: 5 samples at 0.5m resolution will require 30s/m workload, 15 samples at 2m resolution will need 12.5s/m to be deployed, and 25 samples at 4m resolution will require 8.8s/m workload. Thus, the least intensive workload yielding 0.75m accuracy will be achieved with the last combination.

4.4 Prediction on Deployed Edges

To validate the proposed approach we apply our method on 3 edges in two real world navigation fields, one on a university campus (first two edges) and one in a office building environment (remaining edge). The first edge is 15m long, marked every 4m, using 30 samples per point and 1 beacon every 4m. The second edge is 16m long, marked every 16m, using 20 samples per point and 1 beacon every 8m, while the third edge is 21m long, marked every 1m, using 10 samples per point and 1 beacon every 4m.

In Figure 10 we show the predicted and the actual accuracy measured for all three edges. The difference between predicted and actual accuracy is highest on the third edge for the proposed approach. The predicted accuracy was 0.83m while the actual accuracy measured was 0.69m, that is 0.14m difference, corresponding to 16%. For the k-NN approach, the highest difference was 0.15m, with predicted error of 2.36m and measured error of 2.21m (that is 7%) on the first edge.

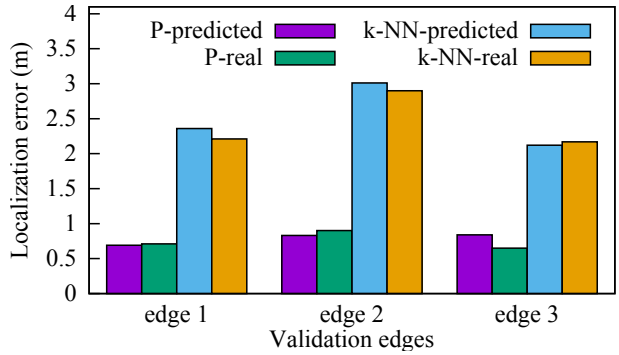


Figure 10: Predicted and achieved accuracy of the k-NN baseline and the proposed approach P on the 3 evaluation edges.

We applied our method to a real world navigation field deployed on a university campus (Described in the following section). The navigation field has a consistent density of 1 beacon every 3m and 30 samples per point were collected at a 1m resolution. Under this assumption, the predicted localization accuracy is 0.79m for the proposed localization technique, and 2.13m for the k-NN regression approach used as baseline. In the following we show how this difference in accuracy impacts navigation with 6 test subjects with visual impairments.

ID	Gender	Age	Vision Type	Navigation Aid
P1	M	43	Totally blind, born sighted and progressively lost vision until 18	White cane (primary) and guide dog
P2	F	73	Legally blind from age 6, some light sensitivity	White cane
P3	F	35	Totally blind from birth	Guide dog (primary) and white cane
P4	F	67	Totally blind from birth	White cane
P5	M	62	Totally blind from birth	White cane
P6	M	66	Totally blind from age 6, some light sensitivity	White cane

Table 1: Demographic information of our participants.

5. EVALUATION WITH USERS

Testing the measured accuracy of our system quantitatively was important to understand it's limitations, but we also wanted to understand how the localization accuracy achieved with the proposed method would impact the usability for people with visual impairments while navigating. Our localization method was integrated into a smartphone navigation application and tested on a university campus in two separate evaluation sessions. We recruited 6 participants from the surrounding area who were legally blind (see Table 1). While the average age of our participants (avg: 57.6 years, std: 13.8) is representative of the visually impaired population [Leonard], in the future we will expand the evaluation to include youth participants.

We hypothesized that the improved accuracy of the localization method proposed in the previous sections would help participants make less errors near turns. We also hypothesized they would respond more positively in interviews after using the proposed localization method due to the fewer errors encountered.

The environment used in the study spanned several buildings and the university quad area (outdoors). The total length of all deployed paths is 530m across 21 edges. Session 1 was conducted in September 2015, and the route (Route 1) included 7 indoor and 5 outdoor edges, totaling 12 edges. Session 2 occurred six months later (March 2016) and only included indoor areas (Route 2) due to inclement weather conditions. The total number of edges in the second study was 10. In terms of length, Route 1 was 390m because of the long outdoor edges, while Route 2 was shorter, totaling 230m (see Figure 2). Nonetheless, both edges had a similar complexity and shared 6 common segments.

We requested that the participants navigate the designated routes with whatever navigation aid they typically used. All participants used a white cane, except P3 who navigated both routes with her guide dog. During both sessions, one experimenter always closely followed the participant to ensure that there were no safety issues while another experimenter recorded videos of the experiments. After completing the route, the participants were escorted back to a conference room for an interview regarding their experience navigating.

5.1 Session 1: K-NN Baseline Localization

In this first session, the navigation application relied on the k-NN localization method, described previously in [Ahmetovic et al. 2016b], to estimate the user's position. The participants kept the phone in a belt pouch and only interacted with the interface through a wrist-worn bluetooth button that allowed them to repeat previous instructions.

Each participant used the navigation application to travel from one building in the university's computer science department to the university bookstore on a 390m long indoor/outdoor route. The route traversed 2 buildings, a set of stairs, 2 bridges, and an outdoor quad area.

5.2 Session 2: Improved Localization

The second session occurred six months after the first, and all but one participant (P1) returned. For this trial, the navigation application was improved to have all functionality included on the smartphone's screen (via iOS VoiceOver), so participants held the phone while navigating. Additionally, the localization method used was the localization technique proposed in this contribution, which relies on PDR and particle filtering for localization.

Due to winter weather conditions, the route followed in the previous experiment was deemed dangerous. Thus, we modified the previous route to go from a university library to a cafe through three buildings, an elevator, and two bridges. The new route was 230m long, and it included 6 common segments with Route 1.

5.3 Analysis of Missed Turns

To measure if the proposed localization method improves the navigation for participants in a quantifiable manner, we compared the number of missed turns made by each participant in Session 1 and Session 2. We defined a missed turn as any time the participant tried to turn either too early or too late and missed the correct turn area. Two members of the research team analyzed the videos from the sessions to record these missed turn events per each route segment.

Because Route 1 and Route 2 differed, we separated the missed turn events in each session into shared and non-shared sections of the routes. For the sections of route that overlapped, we performed a paired t-test. Results show (See Figure 11) that there was a significant difference in the number of missed turns with k-NN localization ($M = 0.467$, $SD = 0.507$) and the proposed method ($M = 0.267$, $SD = 0.450$); $t(29) = -1.795$, $p = 0.042$.

An unpaired t-test was also performed for the sections of route that did not overlap (but shared similar characteristics such as navigation complexity). For these segments, there was also a significant difference in the number of missed turns between the localization method used in Session 1 ($M = 0.35$, $SD = 0.489$) and Session 2 ($M = 0$, $SD = 0$); $t(19) = -3.199$, $p = 0.002$ (See Figure 11). In both cases, participants missed fewer turns while using the more accurate proposed localization method.

5.4 Interview Feedback

The interviews with participants in both sessions were mostly positive as all the participants appreciated the potential of the proposed system to improve the independence of individuals with visual impairments during mobility. However, in Session 1 we received a lot of critical feedback about the timing of instructions. P6 walked very quickly, so the application often provided the instructions as he passed intersections.

> The directions were accurate, I would just hope they would be quicker when you are going to make your turn. – P6

This lack of advance warning when using the k-NN method was

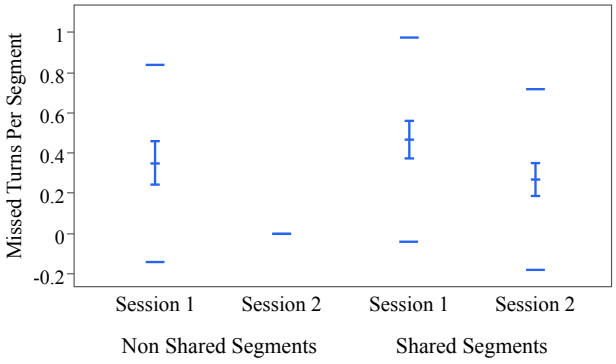


Figure 11: Missed turns per segment in each session, separated by the segments that were shared or not shared between the two routes. The middle blue line is the mean, the shorter arm is the std. error, and the extreme blue lines denote the std. dev.

largely due to the inaccuracy of the localization when the user came to a turn. Participants wanted the timing of these commands to either be adjustable or take into account the user's speed. P3, a guide dog user, needed to have commands sufficiently in advance to instruct their dog:

> Because it doesn't tell me turn right or left until I'm right on it, it doesn't allow me to give my dog the heads up that we are going to turn right or left – P3

This complaint did not come up agian in the second session, as the proposed localization method was more accurate and accounted for the user's speed in the motion model.

> I think it predicted better based on my pace when I was going to be approaching the turn. I didn't feel like I was going to blow past the turn and have no idea when it would say 'approaching' [. . .] I felt like I could rely on it more. – P3

P6 was surprised in Session 2 because he was expecting the application to be more inaccurate based on his experience in the previous session.

> I wasn't trusting it as much as I should have. When it said 'turn right', that's turn right! Yes, it is a lot quicker. – P6

Other participants also reported that the application with the improved localization method was more accurate than the previous version using k-NN localization.

> Oh, it was very easy. – P5

> I can't remember the last one, but when it said 'approaching' I felt like I knew that [the turn] really was coming up soon. [. . .] I felt like 'Ha, I could almost do this on roller skates!' – P4

The participants mentioned some suggestions for the navigation interface. P6 suggested we add more sounds to indicate the turn locations.

> [W]hen you hear 'turn right' and you hear a beep, you know, your body - it's so easy to react, it's so much easier to react to just a quick sound. – P6

P5 and P6 both wished the system was more aware of their position relative to the walls of the hallway, as they wanted warnings if they would hit a plant, pole, or overhang on one side. NavCog is only aware of the user's position down the length of the hallway, so precise alerts like that are not yet possible with the system. Participants said general alerts such as *"There is a staircase on the right side of the hallway"* may be sufficient.

The participants reccomended additional features and modes for NavCog during the interviews, such as a preview mode and rerouting due to accessibility issues. We think these are important features for a navigation application, and we hope to implement them in the future.

6. CONCLUSION AND FUTURE WORK

In order to achieve robust turn-by-turn navigation assistance for people with visual impairments, we improved the localization capabilities of the NavCog system [Ahmetovic et al. 2016a, Ahmetovic et al. 2016b] to enable a sufficient level of accuracy while containing the costs required for defining new navigation fields. The proposed localization technique models the navigation field as a graph of one-dimensional line segments and introduces two new features: 1) BLE beacons RSSI probability distribution estimation, and 2) pedestrian dead reckoning (PDR).

The improvements to the NavCog system allowed us to achieve, in our test environment, a localization accuracy of 0.68m on average with a smaller number of beacons (up to a minimum of 1 beacon every 6m), and a limited installation workload. The accuracy achieved by the improved method is also better than the previously used k-nearest neighbor regression baseline, proposed in [Ahmetovic et al. 2016b], which reached an average accuracy of 1.2m during our evaluation.

We also proposed a design method for defining new navigation fields to provide the desired localization accuracy level while minimizing the number of beacons required and the installation workload. We achieve this goal by studying how beacon density, fingerprinting resolution, and the number of signal fingerprints influence the localization accuracy. We validate the proposed approach on 3 different paths in 2 testing environments. The localization accuracy on these paths predicted with our method is consistent with the measured localization accuracy. As a future work, we will also evaluate the impact of the interference caused by passer-by and objects on the localization accuracy achieved by our method.

We evaluate the impact of the increased localization accuracy on the capability to assist users in navigating through unfamiliar environments with a turn-by-turn navigation procedure. The evaluations with 6 blind test subjects were performed across the two routes, 230m and 390m long, that involved 12 and 14 decision points respectively. At each decision point the participants performed one of the following activities: turning, finding a door, boarding an elevator.

The second experiment was performed at a distance of 6 months to limit the effect of learning on the navigation capability of test subjects. The results of the user evaluation show that our technique is capable of accurately guiding individuals with visual impairments through a testing navigation field unfamiliar to the users. Current tests in two different navigation fields, deployed inside an office building environment and a shopping mall, are showing analogous promising results.

7. ACKNOWLEDGMENTS

The research on the NavCog navigation assistant was partially supported by Shimizu Corporation.

8. REFERENCES

[Ahmetovic et al. 2014] Dragan Ahmetovic, Cristian Bernareggi, Andrea Gerino, and Sergio Mascetti. 2014. Zebrarecognizer: Efficient and precise localization of pedestrian crossings. In *2014 22nd International Conference on Pattern Recognition (ICPR)*. IEEE, 2566–2571.

[Ahmetovic et al. 2016a] Dragan Ahmetovic, Cole Gleason, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016a. NavCog: turn-by-turn smartphone navigation assistant for people with visual impairments or blindness. In *Web for All Conference*. ACM.

[Ahmetovic et al. 2016b] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016b. NavCog: A Navigational Cognitive Assistant for the Blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM.

[Ahmetovic et al. 2016] Dragan Ahmetovic, Roberto Manduchi, James Coughlan, and Sergio Mascetti. 2016. Mind your crossings: Mining GIS imagery for crosswalk localization. *ACM Transactions on Accessible Computing* (2016).

[Ahmetovic et al. 2015] Dragan Ahmetovic, Roberto Manduchi, James M Coughlan, and Sergio Mascetti. 2015. Zebra crossing spotter: automatic population of spatial databases for increased safety of blind travelers. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. ACM, 251–258.

[Amemiya et al. 2004] 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 *Virtual Reality, 2004. Proceedings. IEEE*. IEEE, 165–287.

[Benjamin and Malvern 1973] JM Benjamin and J Malvern. 1973. The new C-5 laser cane for the blind. In *Proc. Carnahan Conf. on Electronic Prosthetics*. 77–82.

[Blasch et al. 1996] 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.

[Brajdic and Harle 2013] Agata Brajdic and Robert Harle. 2013. Walk detection and step counting on unconstrained smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 225–234.

[Brock and Kristensson 2013] Michael Brock and Per Ola Kristensson. 2013. Supporting blind navigation using depth sensing and sonification. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 255–258.

[Chumkamon et al. 2008] Sakmongkon Chumkamon, Peranitti Tuvaphanthaphiphat, and Phongsak Keeratiwintakorn. 2008. A blind navigation system using RFID for indoor environments. In *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on*, Vol. 2. IEEE, 765–768.

[Coughlan and Yuille 2000] James M Coughlan and Alan L Yuille. 2000. The Manhattan world assumption: Regularities in scene statistics which enable Bayesian inference. In *NIPS*, Vol. 2. 3.

[Dockstader and Tekalp 2001] Shiloh L Dockstader and A Murat Tekalp. 2001. Multiple camera tracking of interacting and occluded human motion. *Proc. IEEE* 89, 10 (2001), 1441–1455.

[Fallah et al. 2012] 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 *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 425–432.

[Faragher and Harle 2015] Ramsey Faragher and Robert Harle. 2015. Location Fingerprinting With Bluetooth Low Energy Beacons. *Selected Areas in Communications, IEEE Journal on* 33, 11 (2015), 2418–2428.

[Faria et al. 2010] José Faria, Sérgio Lopes, Hugo Fernandes, Paulo Martins, and João Barroso. 2010. Electronic white cane for blind people navigation assistance. In *World Automation Congress*. IEEE.

[Fiannaca et al. 2014] Alexander Fiannaca, Ilias Apostolopoulous, and Eelke Folmer. 2014. Headlock: A wearable navigation aid that helps blind cane users traverse large open spaces. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, 19–26.

[Fod et al. 2002] Ajo Fod, Andrew Howard, and MAJ Mataric. 2002. A laser-based people tracker. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, Vol. 3. IEEE, 3024–3029.

[Gonzalez et al. 2007] J Gonzalez, JL Blanco, C Galindo, A Ortiz-de Galisteo, JA Fernández-Madrigal, FA Moreno, and JL Martinez. 2007. Combination of UWB and GPS for indoor-outdoor vehicle localization. In *Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on*. IEEE.

[Hilsenbeck et al. 2014] Sebastian Hilsenbeck, Dmytro Bobkov, Georg Schroth, Robert Huitl, and Eckehard Steinbach. 2014. Graph-based data fusion of pedometer and WiFi measurements for mobile indoor positioning. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 147–158.

[Iwahashi 1983] Hideyuki Iwahashi. 1983. *Toward white wave - Story of Seiichi Miyake (in Japanese)*. Traffic Safety Research Center.

[Kacorri et al. 2016] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Dragan Ahmetovic, Hironobu Takagi, and Chieko Asakawa. 2016. Supporting Orientation of People with Visual Impairment: Analysis of Large Scale Usage Data. In *International ACM SIGACCESS Conference on Computers and Accessibility*. ACM.

[Kleege 2006] Georgina Kleege. 2006. Visible Braille/invisible blindness. *Journal of visual culture* 5, 2 (2006), 209–218.

[Legge et al. 2013] Gordon E Legge, Paul J Beckmann, Bosco S Tjan, Gary Havey, Kevin Kramer, David Rolkosky, Rachel Gage, Muzi Chen, Sravan Puchakayala, and Aravindhan Rangarajan. 2013. Indoor navigation by people with visual impairment using a digital sign system. *PloS one* 8, 10 (2013), e76783.

[Leonard] Robin McInerney Leonard. *Statistics on vision impairment: A resource manual*.

[Manduchi et al. 2013] Roberto Manduchi, James Coughlan, Klaus Miesenberger, James M Coughlan, and Huiying Shen. 2013. Crosswatch: a system for providing guidance to visually impaired travelers at traffic intersection. *Journal of Assistive Technologies* 7, 2 (2013), 131–142.

[Manduchi and Kurniawan 2012] Roberto Manduchi and Sri

Kurniawan. 2012. *Assistive technology for blindness and low vision*. CRC Press.

[Muja and Lowe 2009] Marius Muja and David G Lowe. 2009. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *VISAPP (1)* 2 (2009), 331–340.

[Murphy 2012] Kevin P Murphy. 2012. *Machine learning: a probabilistic perspective*. MIT press.

[Nakajima and Haruyama 2012] Madoka Nakajima and Shinichiro Haruyama. 2012. Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing. In *Communications in China*. IEEE.

[Peraković et al. 2015] Dragan Peraković, Marko Periša, and Vladimir Remenar. 2015. Model of guidance for visually impaired persons in the traffic network. *Transportation research part F: traffic psychology and behaviour* 31 (2015), 1–11.

[Poulsen 1982] Torben Poulsen. 1982. Acoustic traffic signal for blind pedestrians. *Applied Acoustics* 15, 5 (1982), 363–376.

[Pressey 1977] N Pressey. 1977. Mowat sensor. *Focus* 11, 3 (1977), 35–39.

[Ryckaert et al. 2004] Julien Ryckaert, Philippe De Doncker, René Meys, Arnaud de Le Hoye, and Sophie Donnay. 2004. Channel model for wireless communication around human body. *Electronics letters* 40, 9 (2004), 543–544.

[Saito et al. 2007] Shigeru Saito, Atsushi Hiyama, Tomohiro Tanikawa, and Michitaka Hirose. 2007. Indoor marker-based localization using coded seamless pattern for interior decoration. In *2007 IEEE Virtual Reality Conference*. IEEE.

[Williams et al. 2014] Michele A Williams, Caroline Galbraith, Shaun K Kane, and Amy Hurst. 2014. just let the cane hit it: how the blind and sighted see navigation differently. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. ACM, 217–224.