Assistive Technology
The Official Journal of RESNA

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/uaty20

Evaluating the efficacy of UNav: A computer vision-based navigation aid for persons with blindness or low vision

Anbang Yang, Nattachart Tamkittikhun, Giles Hamilton-Fletcher, Vinay Ramdhanie, Thu Vu, Mahya Beheshti, Todd Hudson, Wachara Riewpaiboon, Pattanasak Mongkolwat, Chen Feng & John-Ross Rizzo

To cite this article: Anbang Yang, Nattachart Tamkittikhun, Giles Hamilton-Fletcher, Vinay Ramdhanie, Thu Vu, Mahya Beheshti, Todd Hudson, Wachara Riewpaiboon, Pattanasak Mongkolwat, Chen Feng & John-Ross Rizzo (13 Aug 2024): Evaluating the efficacy of UNav: A computer vision-based navigation aid for persons with blindness or low vision, Assistive Technology, DOI: 10.1080/10400435.2024.2382113

To link to this article: https://doi.org/10.1080/10400435.2024.2382113

Published online: 13 Aug 2024.

ASSISTIVE TECHNOLOGY
https://doi.org/10.1080/10400435.2024.2382113

Taylor & Francis
Taylor & Francis Group

Evaluating the efficacy of UNav: A computer vision-based navigation aid for persons with blindness or low vision

Anbang Yang, MSc[a], Nattachart Tamkittikhun, MSc[b], Giles Hamilton-Fletcher, PhD[c,d], Vinay Ramdhanie, BSc[e], Thu Vu, BSc[f], Mahya Beheshti, MD[c], Todd Hudson, PhD[c], Wachara Riewpaiboon, MD[g], Pattanasak Mongkolwat, PhD[b], Chen Feng, PhD[a], and John-Ross Rizzo, MD, MSCI[a,c,e,g]

[a]Department of Mechanical and Aerospace Engineering, NYU Tandon School of Engineering, Brooklyn, New York, USA; [b]Faculty of Information and Communication Technology, Mahidol University, Nakhon Pathom, Thailand; [c]Department of Rehabilitation Medicine, NYU Grossman School of Medicine, New York, New York, USA; [d]Department of Ophthalmology, NYU Grossman School of Medicine, New York, New York, USA; [e]Department of Biomedical Engineering, NYU Tandon School of Engineering, Brooklyn, New York, USA; [f]Department of Computer Science and Engineering, NYU Tandon School of Engineering, Brooklyn, New York, USA; [g]Ratchasuda Institute, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Nakhon Pathom, Thailand

ABSTRACT

UNav is a computer-vision-based localization and navigation aid that provides step-by-step route instructions to reach selected destinations without any infrastructure in both indoor and outdoor environments. Despite the initial literature highlighting UNav's potential, clinical efficacy has not yet been rigorously evaluated. Herein, we assess UNav against standard in-person travel directions (SIPTD) for persons with blindness or low vision (PBLV) in an ecologically valid environment using a non-inferiority design. Twenty BLV subjects (age = 38 ± 8.4; nine females) were recruited and asked to navigate to a variety of destinations, over short-range distances (<200 m), in unfamiliar spaces, using either UNav or SIPTD. Navigation performance was assessed with nine dependent variables to assess travel confidence, as well as spatial and temporal performances, including path efficiency, total time, and wrong turns. The results suggest that UNav is not only non-inferior to the standard-of-care in wayfinding (SIPTD) but also superior on 8 out of 9 metrics, as compared to SIPTD. This study highlights the range of benefits computer vision-based aids provide to PBLV in short-range navigation and provides key insights into how users benefit from this systematic form of computer-aided guidance, demonstrating transformative promise for educational attainment, gainful employment, and recreational participation.

ARTICLE HISTORY
Accepted 9 July 2024

KEYWORDS
assistive technology; indoor navigation; outdoor navigation; UNav; visually impaired

Introduction

Around 295 million people globally are visually impaired, with 43 million being blind. By 2050, global estimates are projected to increase to 474 million for visual impairment and 61 million for blindness (Kruk & Pate, 2020). Visual impairments very often significantly hinder spatial cognition, situational awareness, and equity (Giudice, 2018; Pascolini & Mariotti, 2012), negatively impacting their ability to navigate a wide range of scenarios (Hakobyan et al., 2013; Rizzo et al., 2023). Prior studies exploring navigation in unfamiliar environments for people who are blind or have low vision (PBLV) have found that 70% of the interviewees had not visited key locations because of physical challenges and the need for sighted guides (Jeamwatthanachai et al., 2019a). If PBLV did visit new areas, a sighted guide was the primary way they built up their initial mental map or spatial cognition of an area. After this process, when it comes to locating themselves within a space, 30% primarily use landmarks, 24% use environmental cues (e.g. sources of light, sound, and odor), 9% use the floor texture, and 34% primarily use a sighted guide. But when it came to physically navigating to destinations, both floor texture (16%) and usage of sighted guides (56%) were increasingly relied upon (Jeamwatthanachai et al., 2019b).

Since the 1960s, technological advances have produced a range of assistive technologies or tools for PBLV (Han et al., 2023; Kandalan & Namuduri, 2019). Navigation tools vary greatly in which kinds of support they provide (Ball, 2008) – from basic obstacle avoidance with electronic travel aids (ETAs) all the way to user localization and wayfinding support with electronic orientation aids (EOAs) (Goswami et al., 2023; Xu et al., 2023). Many of these come with challenges including high cost, limited availability, real-world inaccuracies, over-reliance on extensive infrastructure, dependencies on comprehensive training and poor ergonomics, poor esthetics, and weak usability characteristics (dos Santos et al., 2022; Hersh, 2022; McGrath & Astell, 2017; Okonji & Ogwezzy, 2019; Senjam et al., 2019). Some smartphone apps such as GoodMaps (GoodMaps, 2024) require areas to be pre-mapped using expensive LiDAR rigs in order to support indoor navigation. While ASSIST (Nair et al., 2022), and NavCog (Ahmetovic et al., 2016; Sato et al., 2019) use BLE-beacon-based infrastructure within the environment for indoor navigation. Both approaches create difficulties, as the required materials may be difficult to obtain or use at scale with adequate coverage. Some indoor navigation

CONTACT Chen Feng cfeng@nyu.edu; John-Ross Rizzo johnross.rizzo@nyulangone.org Department of Mechanical and Aerospace Engineering, NYU Tandon School of Engineering, Brooklyn, NY 11201.

aids like MagNav can utilize the preexisting geomagnetic field maps sensed through the physical infrastructure of indoor environments to infer position, using changes in magnetometer readings alongside accelerometers, gyroscopes, and pedometers (Giudice et al., 2019). Some systems also have the user contributes to the smooth running of the navigation system, for example, having the user manually confirm when specific places are reached in the journey, to prompt the next set of instructions with Navatar (Fallah et al., 2012). Other solutions have users initially map the route in augmented reality themselves to explore later or share it with others using the Clew app (Yoon et al., 2019).

Issues surrounding the accessibility and availability of EOAs are further exacerbated in low- and middle-income countries (LMICs), such as Thailand. LMICs often suffer from over-representations of blindness and low vision in their populations and also more inaccessibility in the built environment in both urban and rural settings, particularly in healthcare and educational settings (Bourne et al., 2003; Gilbert & Foster, 1993; Isipradit et al., 2014; Jenchitr et al., 2011; Prabhasawat et al., 2007). Thailand has sought to enhance medical rehabilitation for persons with disabilities through the Thailand Rehabilitation Act of 1991, and its evolution into the Persons with Disability Empowerment Act of 2007. Its 2010 legislative update included a universal health benefit package with orientation and mobility training and white canes for use as a primary mobility tool for PBLV. However, any secondary mobility tools, such as EOAs, need to suit the currently available physical/digital resources in Thailand. Unfortunately, access to the hardware and physical infrastructure requisite for many EOAs is often less available and more expensive when considering local buying power as well as running the risk of interoperability bottlenecks (Bohonos et al., 2007; Suebsin & Gerdsri, 2012). As a result, solutions that do not require hardware infrastructure may provide the most affordable, accessible, and scalable options for assisting PBLV globally. Fortunately, as part of Thailand's push toward the "digitization of society," cities such as Bangkok feature excellent network connectivity, with widespread 5 G availability (nPerf, 2024; Statista, 2024a; 2024b). The high speed of 5 G data transfer (~140Mbps) allows for fast high-quality uploads and downloads from the users' smartphone or EOA. This means that any intensive data processing required for navigation assistance can be pushed or offloaded to edge-processors or server-infrastructure and rapidly communicated back to the user. Overall, this means that utilizing Bangkok's existing network connectivity for enhancing navigation is a viable solution for providing cutting-edge EOAs at lower cost and higher availability.

To address these challenges, we created a novel computer vision-based navigation aid, called UNav (Yang et al., 2022). UNav only requires standard RGB images from the user, on either a mobile application or a smart wearable with integrated hands-free cameras, and applies visual place recognition (VPR) (Sheng et al., 2021) techniques to localize them (Yang et al., 2023). VPR affords the ability to efficiently reference the users' local image-based visual features (extracted from the image frames of the app or wearable) against environmental visual features which have been mapped to a floor plan for georeferencing (Pan et al., 2022; Sarlin et al., 2019). This provides the users' location and orientation, which when combined with a specified destination and a path planner algorithm, allows UNav to provide step-by-step wayfinding instructions to the user. This calculation is typically executed on a remote server to localize users in near real-time (Yang et al., 2023). As a result, effective use in real-world scenarios requires accounting for variable or poor network conditions. For this, we use rate-adaptation algorithms such as REBERA (Azzino et al., 2023), to ensure that high-quality images are uploaded from the user to the server even under poor network conditions. This enhances performance and supports more seamless remote processing on servers, boosting the performance of UNav and potentially a variety of additional remote microservices (e.g. object detection, text recognition, and scene descriptions).

To make navigation aids more accessible and responsive to the unique needs of PBLV, UNav was designed with a low barrier to entry. It requires only standard RGB camera images (rather than RGBD (Goswami et al., 2023)) from a single video walkthrough of the environment alongside a basic two-dimensional floor plan. The time needed for image capture in the field is the same as walking through each path that will be supported, while the tools that combine video and floor plans into topometric maps are semi-automated, taking ~5 minutes to map an area the size of a New York City block (about 264 × 900 feet). This approach is scalable to large environments and critically does not require external infrastructure (e.g. beacons (Cheraghi et al., 2017)) or specialized mapping equipment (e.g. LiDAR scanners (Chen et al., 2023; Jain & Patel, 2023; Otero et al., 2020)). UNav's visual place recognition method is highly accurate, both in localizing the user (spatial error <1 m) and establishing their orientation (<5° error). Furthermore, as new visual data is received during journeys, UNav's representation of space is dynamically updated to reflect any changes in the environment, making this a robust and self-sustaining approach that evolves with a future-proofed motif. This results in a navigation tool that is precise, cost-effective, and adaptive to the environment. Furthermore, additional important environmental details can be added (e.g. room numbers and points-of-interest). This provides a technological foundation to explore how specific types of information can assist the navigational needs of PBLV as well as test its effect on their spatial cognition and memory.

While UNav has been technically validated previously (Yang et al., 2022), its efficacy in supporting navigation for PBLV in real-world scenarios has not yet been evaluated. Here, we conducted a non-inferiority trial, comparing UNav's navigation support against standard in-person travel directions (SIPTD) over short-range distances (<200 m) between key destinations in unfamiliar indoor/outdoor environments. Navigation performance was assessed using both spatial, temporal, and user-confidence metrics. Spatial metrics include navigation success, path efficiency, wrong turn incidences, and cane contact frequencies. Temporal metrics include total navigation time, number of idles (subject comes to a complete stop), idle time (time spent stationary), and gait speed. User confidence was gauged by the number of help requests. Our hypothesis is that UNav would be non-inferior to standard in-person travel directions. We predicted that UNav would match, if not best, the performance of our nine dependent variables for PBLV when traveling in and through unfamiliar, ecologically valid environments.

Methods

The primary aim of this study is to assess and contrast the navigational performance of PBLV using two distinct support mechanisms: the UNav wayfinding system (Yang et al., 2022), which offers computer-based, systematic step-by-step instructions to the user (termed the "active" condition); and conventional in-person instructional guidance from trained chaperones (termed the "passive" condition).

Participants

To test our hypothesis, a cohort of $n = 20$ participants with impairments were recruited (see Table 1). We use visual functions to classify how well each participant can perform visual-related activities (National Research Council, 2002). This consisted of 12 people who are blind (Visual Function [VF] rating = 6), 3 persons with poor to very poor visual conditions (VF = 5 or 4), and 5 persons with fair visual function (VF = 3). All participants are considered legally blind by Thai legislation. This group had a mean age of 38.3 years (±8.4 SD), with 9 females. Table 1 provides an overview of the participants' demographics, including their age, gender, educational level, work status, onset of age of visual impairment, nature of vision loss, prior orientation and mobility training, cane usage, health rating, and visual-function rating.

The study (protocol number MU-CIRB 2021/288.3105) was approved by the Mahidol University Central Institutional Review Board, Certificate of Approval No. MU-CIRB 2021/180.2708.

The recruitment process was conducted by the Thai Graduate Society of the Blind and the Visual Impairment Section at Ratchasuda Institute Faculty of Medicine Ramathibodi Hospital, Mahidol University. The snowball sampled people from Bangkok and its perimeter, including Nakhon Pathom, according to the following inclusion and exclusion criteria. To be included as a participant, the individual needed to be older than 18 years and medically diagnosed with visual impairment (legal blindness, including acuity and/or field deficits). He/she had to be currently using a primary assistive device for mobility (e.g. white cane or guide dog) and be able to travel without the aid of another person. However, an individual would be excluded if he/she had significant injury to upper and/or lower extremities. People who had auditory impairments and/or were pregnant could not participate in the project. Recruitment also excluded those with comorbid neurological illnesses and confounding medical conditions. People with somatosensory impairments to the trunk or torso that precludes the use of haptic interfaces would also not be recruited.

Table 1. Participant demographics.

No.	Age	Gender	Education level	Work status	Age onset of VI	Loss of vision	O&M training	Cane use age	Health rating	VF rating
1	54	Male	PhD	Gov employee	0	Gradually	Vocational center	15	3	6
2	38	Male	Master	Gov employee	0	Gradually	Blind school	8	3	5
3	52	Female	PhD	Gov employee	0	Gradually	Ratchasuda College	10	3	6
4	33	Female	Master	Gov employee	0	Gradually	Ratchasuda College	18	3	6
5	52	Male	Master	Gov employee	35	Sudden	Ratchasuda College	35	3	6
6	32	Male	Secondary school	Company employee	0	Gradually	Blind school	12	4	3
7	41	Female	Graduate	Self- employed	34	Gradually	Ratchasuda College	34	4	6
8	33	Male	Secondary school	Company employee	7	Gradually	Vocational center	16	3	3
9	32	Female	Graduate	Company employee	0	Gradually	Blind school	12	3	3
10	42	Male	Secondary school	Company employee	late life	Gradually	Ratchasuda College	35	3	6
11	48	Male	Vocational certificate	Self- employed	late life	Gradually	Ratchasuda College	No	3	3
12	33	Male	Graduate	Self- employed	12	Gradually	Blind school	25	1	4
13	41	Female	Secondary school	Company employee	0	Gradually	Vocational center	7	3	6
14	34	Male	Graduate	Company employee	0	Gradually	Blind school	15	3	6
15	45	Female	Graduate	Company employee	0	Never seen	Blind school	10	3	6
16	37	Female	Graduate	Company employee	0	Gradually	Blind school	10	2	6
17	28	Female	Graduate	Self- employed	0	Gradually	Special education center	20	2	3
18	35	Male	Secondary school	Self- employed	0	Gradually	Blind school	9	1	6
19	23	Female	Secondary school	Self- employed	0	Gradually	Blind school	8	2	4
20	34	Male	Graduate	Self- employed	0	Never seen	Blind school	7	3	6

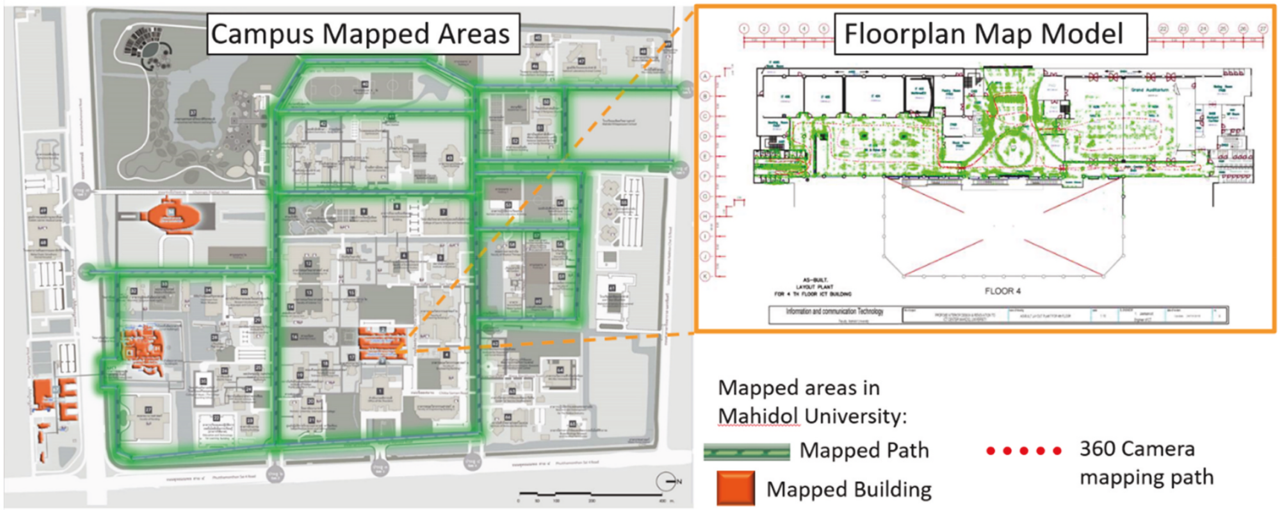

Figure 1. Mapped areas. Left image shows outside areas of Mahidol Campus that have been mapped for the UNav system (in green) as well as indoors for specific buildings (in orange). Right image shows an example of an indoor mapped area. This shows the visual features of the environment (green dots) that are mapped onto a floor plan, with the mappers pathing shown in red.

Recruited participants were introduced with brief information about the research project by staff from the Thai Graduate Society of the Blind and the Visual Impairment Section at Ratchasuda Institute and then referred to the researcher team. The team provided detailed information to the participants including risks and benefits of the study. Participants were asked general questions about their visual impairment and then asked to sign a consent form to participate in the study.

Setting

The study was conducted in the Salaya campus of Mahidol University in the city of Nakhon Pathom in Thailand, just outside Bangkok. For this, we mapped seven buildings on campus consisting of the Faculty of Information and Communication Technology (ICT), Office of President (OP), Library (Lib), Mahidol Prince Hall (MH), Mahidol Learning Center (MLC), Music College (MS), and Ratchasuda College (RS) – See Figure 1. For this study, we mapped ~92900 m^2 in total (1 million sq ft), with approximately 75% of this being inside buildings at Mahidol University and 25% for the surrounding areas or outside routes between buildings. An area was considered "mapped" when visual features from our 360° camera walkthrough were co-registered to a floor plan or layout (indoor/outdoor), generating a topometric map that acts as a lightweight digital twin of the real environment. Next, destination points were added and labeled in the topometric map. These topometric maps were then used to create 20 indoor routes, which were supplemented with 4 outdoor routes traveling in-between buildings.

Design

This study employs a non-inferiority, repeated-measure crossover design with counterbalancing. The non-inferiority design indicates what constitutes a successful outcome (i.e. that UNav is either equivalent to, or superior to, existing best practices). Furthermore, this design choice does not impede our ability to determine performance "superiority" on measurable outcomes through inferential statistics. The repeated-measure crossover design allows us to evaluate each participant in each condition, increase internal validity, as well as analyze ordering effects. This design also provides the highest available statistical power, as each participant acts as their own control, reducing inter-subject variability in our comparisons.

Apparatus

Mapping

To localize participants and provide step-by-step navigation instructions to them in the active condition, creating an accurate topometric map of the environment is crucial (Mahdi & Xinming, 2022). The Mahidol University maps used for testing were created from visual feature data gathered by a 360-view camera. Data were obtained for each floor using a "zigzag" trajectory, ensuring maximum area coverage (Ramalingam et al., 2022). Three loops were recorded on each floor to ensure map quality and consistency. Post-recording, we aligned the visual feature map with the floor plan as per UNav guidance, creating a topometric map suitable for the path planner algorithm to provide navigation instructions. This results in a digital representation of the environment (free of physical infrastructure), which can be used as a digital twin to explore the real environment with our VIS^{4}ION hardware (Azzino et al., 2023; Gui et al., 2019; Yuan et al., 2022), or by downloading our smartphone app.

Wearable

In the active condition trials, participants were equipped with a wearable backpack system (see Figure 2) comprising a high-resolution 16MP Autofocus USB Camera with a Microphone, featuring a 1/2.8" IMX298 sensor, attached to the strap of the backpack at the chest level for optimal

Figure 2. UNav system components housed in user backpack. Components supporting the UNav navigation system and data acquisition (left image), and how they are worn by the user (right image). Here we showcase the individual system components (orange boxes) and how they are connected to one another (orange lines). Local computational processing and networking is conducted by the Nvidia Orin board, which receives verbal input from the users' Bluetooth mic and images from a strap-mounted camera. This information is sent to a server for processing. Feedback is sent back by the server to the Nvidia Orin board and is then delivered to the user via Bluetooth to bone-conducted hearing.

field-of-view capture. The camera is a Mini UVC USB2.0 4K Video Webcam, which provides high-resolution visual images to the system. Along with the camera, participants wore a backpack containing an NVIDIA Jetson Orin board that receives the camera's video feed, runs the UNav client software, and provides network connectivity. Power for the system was supplied by a MaxOak laptop battery. The backpack weighs 0.7 kg and the integrated system weighs an additional 3.5 kg, inclusive of a large battery to allow multiple, sequential trials without recharging. For audio feedback, participants used a binaural bone-conduction headset equipped with an integrated microphone. This setup not only enabled them to receive verbal instructions from the system but also allowed them to use voice commands to select or change their destination of interest, all while keeping their ears unobstructed to remain aware of ambient

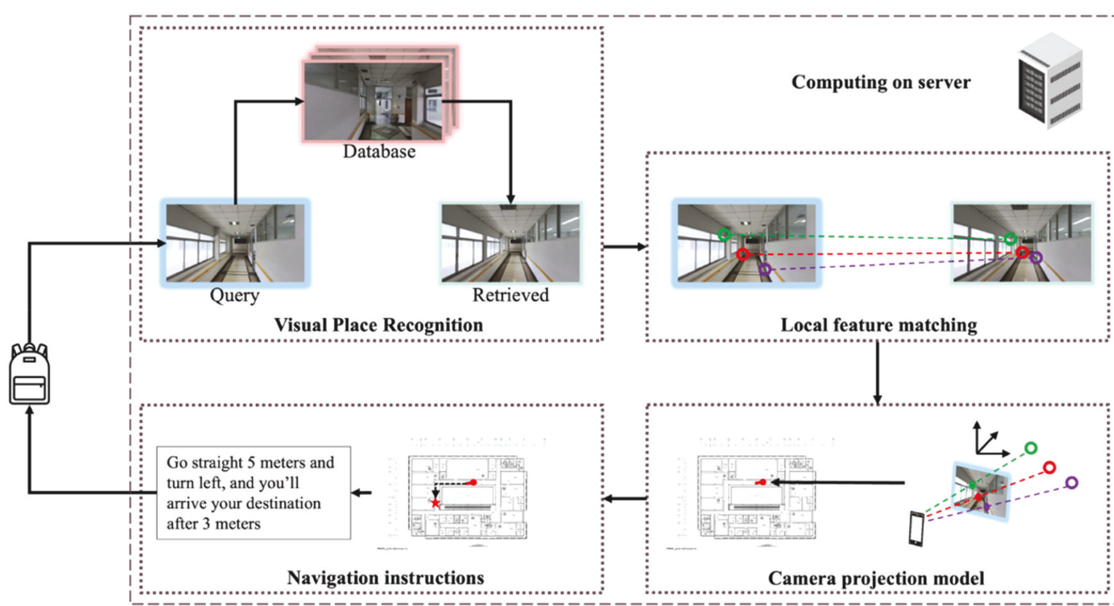

Figure 3. UNav processing architecture. A query image is acquired from a camera mounted on the backpack-based system and is sent to the server. On the server, this query image is referenced against a database of previously acquired images. Using visual place recognition, the most similar image is compared against the query image in terms of local feature matching. The location of these visual features allows the server to calculate the observers' location and orientation on a floor plan using a camera projection model. A route is calculated between the observers' location/orientation and the inputted destination, then a path planner algorithm is used to calculate the next steps necessary to reach the destination, with these instructions sent to the end user via the backpack and Bluetooth headset (not shown).

sound. The participants also wore a 360° camera mounted on a helmet, which was purely used for data collection and not connected to the UNav system.

Server

The server service was run at a node of Nvidia DGX A100 systems at Mahidol AI Center. The service receives the user's selected destination and images captured from the camera of the wearable backpack (see Figure 3). The UNav server software extracts multiple visual features from the client's image and compares this combination of visual features with a database of those gathered from reference images obtained during the environmental mapping phase. All visual features obtained from the 2D images used in the mapping phase are provided with 3D spatial coordinates using a floor plan as well as simultaneous-localization-and-mapping and structure from motion techniques, which creates our 3D topometric digital twin of the environment. The visual features present in reference images that reach a similarity score threshold with the client image (using K-means clustering) are used to calculate the users' position and orientation. Then, a 3D-to-2D "camera projection model" is used to calculate what position and orientation in space the client's perspective must be to yield the spatial arrangement of visual features observed in the client image. Now that the server knows the client's position and orientation, their destination, and the topometric map of the environment, a path planner algorithm provides the route, for which next steps are provided as short sets of instructions. These instructions are then sent back to the wearable backpack to be converted into voice instructions for the user (see Figure 4).

For the server, the UNav process averages 0.62% CPU usage, 13.85% GPU usage, and 4.7% GPU memory for a single user. This means on average that one GPU on one server can support approximately seven users' inquiries simultaneously without a drop in performance. The network bandwidth used is minimal, with sending the image from the client to the server at 70kb, place selection message at 1.4kb, destination selection message at 4.9kb, and server-to-client instructions at 122b. In terms of round-trip time (client-to-server-to-client), navigation instructions are provided back to the client on average at 3.62 seconds (ranging from 0.24 to 5.96 seconds). We have subsequently worked to further reduce this round-trip time using a variety of methods (e.g. parallel processing on GPU/CPUs and use of smaller "sub-maps" to speed up localization processing).

User data

Both active and passive condition trials utilized a GoPro 360 cameras mounted on participants' helmets to capture panoramic motion. This setup captured data related to participants' kinematics and overall navigation performance. Once participants were outfitted with the wearable system and familiarized with its feedback, they were instructed to navigate through predetermined locations on the Salaya campus under both active (receiving instructions from UNav) and passive (receiving instructions from chaperones) support conditions. Each navigation task was designed to be completed within 1–3 minutes and less than 200 m in total distance, i.e. short-range navigation. Tasks were also designed to be confined within a single floor plan to ensure uniformity and consistency in data collection across all trials.

Testing procedures

Participants were first introduced to the UNav system's form, functionality, and how to use it in walking trials. Users were taught how to select destinations via voice instructions using the microphone. Users were also familiarized with how UNav's verbal feedback would communicate direction and distance. Finally, it was ensured that the system's camera was directed forward appropriately.

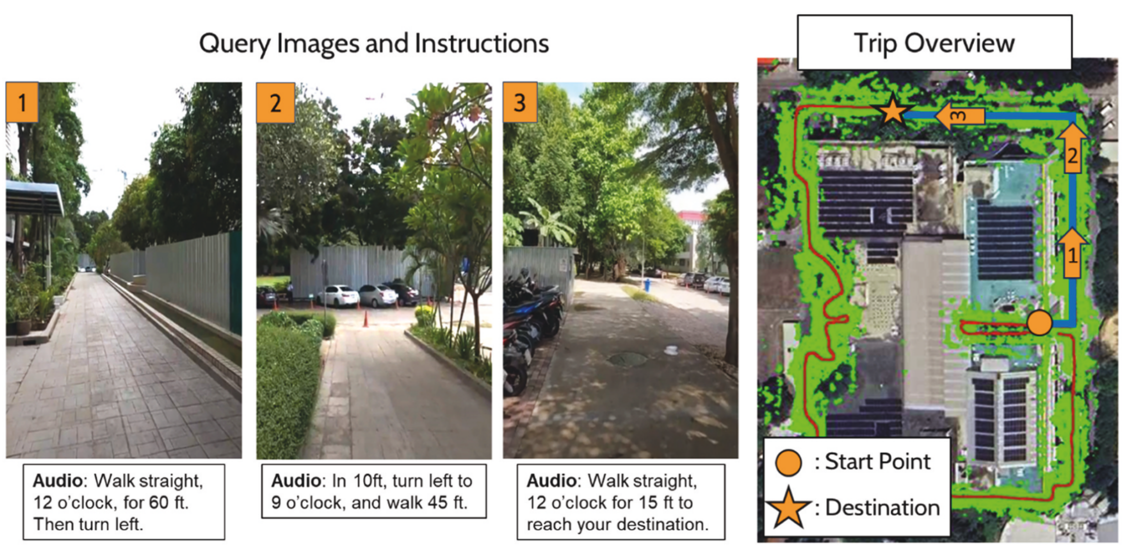

Figure 4. User experience during the active condition (UNav). Query images and instructions show three instances of the users' perspective and subsequent audio instructions to the user from the UNav system during travel between start and end points (right image, blue route). The three instances occur at three different locations during the route (labeled as 1, 2, & 3). Trip overview map also shows visual feature locations (green dots) used to localize the user at any location on the red route.

We assessed navigation across 24 routes, with 4 being outdoors and 20 being indoors. Subjects were encouraged to take part in up to four routes; however, constraints surrounding subjects' participation time and energy as well as data acquisition issues resulted in fewer usable trials. As a result, the data consist of the average metrics for: two participants walking four routes twice (once for passive, once for active condition); four participants walking three routes twice; seven participants walking two routes twice; and six participants walking one route twice. One participant was omitted due to data acquisition issues.

The chosen routes were refined from a larger list based on safety and complexity, i.e. routes must be free of significant hazards (e.g. construction materials), include 2–4 turns to provide a consistent level of complexity, and the total distance being between 50 and 200 meters. These walking trials were across various floors of diverse buildings located on the Salaya campus. The routes were also unfamiliar to participants. This ensured that all participants' prior knowledge was similar and that they had to rely on the intervention for navigation support.

To evaluate the efficacy of the UNav system, each participant navigated the same route twice: once using the UNav's navigation support and once using SIPTD. The order in which participants used the active and passive modes was randomized across subjects to eliminate overall biases due to route familiarity on the 2^{nd} trial. The sequence of conditions was counterbalanced to eliminate potential order effects, with ample time allocated between trials to prevent fatigue and learning effects from over-influencing the outcomes of the study. Of the usable trials, 15 started with active mode followed by passive mode, and 25 were vice versa.

Under passive conditions, participants were set up at the starting point of the selected route. Then, the chaperone instructed them how to get to the destination and let them walk according to their memory. Instructions were given based on body-relative directions and distances, for reaching each turning point and their destination (e.g. "Walk forward for 15 meters, then turn left and walk forward for 10 meters, then turn right and walk forward for 20 meters to reach your destination"). In the active condition, the participants were introduced to the backpack in terms of its physical components and how to use it in the walking trials. They were taught how to issue voice instructions via the microphone on the headset to select the destination. The format of voice navigation instructions they could hear from the headset was also explained, so they were prepared to estimate the distance and direction according to the instructions. When they understood how to use the backpack, the chaperones let them wear it, as well as the headset. Then, the chaperones adjusted the camera on a backpack strap to be level. The participants were brought to the starting point of the route and prompted to select the destination of the route via a voice command. When the system localized participants' current position, the participants would hear the voice navigation instructions and were asked to follow the suggested path while continuing use of their primary mobility aid (e.g. white cane).

In both passive and active conditions, chaperones accompanied the participants and maintained a 1-meter distance from the end user. Chaperones were prepared to answer participants' questions in case they were confused with the navigation instructions from the headset or verbal instructions. The chaperone also informed that the participants if they headed in the wrong direction for more than 3 meters or were close to harmful obstacles for their safety and that they were available for any help requests during the trials. Another chaperone observed the actual path the participants walked and recorded trajectory specifics. In addition, the participants were also asked to wear a helmet with a 360-view camera. This camera did not provide information to the system but was used to collect ground-truth videos from each trial.

Analytic plan

Navigation metrics with a continuous dependent variable and a normal distribution for both conditions of the trial data (using the Shapiro – Wilk test) are analyzed using either paired sample t-tests (e.g. order effects) or between-groups t-tests (e.g. visual function effects). Data that are non-continuous, or feature a group that is not normally distributed, are analyzed using Wilcoxon matched-pairs signed rank tests.

Results

The results indicate that UNav not only matched the performance of standard in-person travel directions (SIPTD), which indicates non-inferiority, but also actually had superior performance in 8 of 9 navigation metrics (with 7 of 9 to a statistically significant degree), including navigation success, travel time, idle period, gait speed, etc. In the following subsections, we will review navigational performance with a focus on spatial performance, temporal performance, and user confidence.

Spatial metrics

Navigation success

Navigation success rates were markedly higher across both conditions. While the passive condition exhibited a 97.5% success rate, the active condition achieved a perfect success rate of 100%. The difference, although noteworthy, was not suitable for statistical significance testing due to the ceiling effect present in the data.

Path efficiency

For each subject, we compared the ratio of the actual distance traveled against the optimal trajectory, in which lower values indicate more efficient paths. Here, we found that path efficiency was significantly better in the active condition (M = 1.17) than in the passive condition ($M = 1.38$), z = −3.59, $p < 0.001$, effect size (z/(√N) (Pallant, Julie and Manual, 2011)) = 0.85 (see Figure 5). This suggests that navigation with the UNav system is more efficient, aiding users in selecting more direct routes to their destinations.

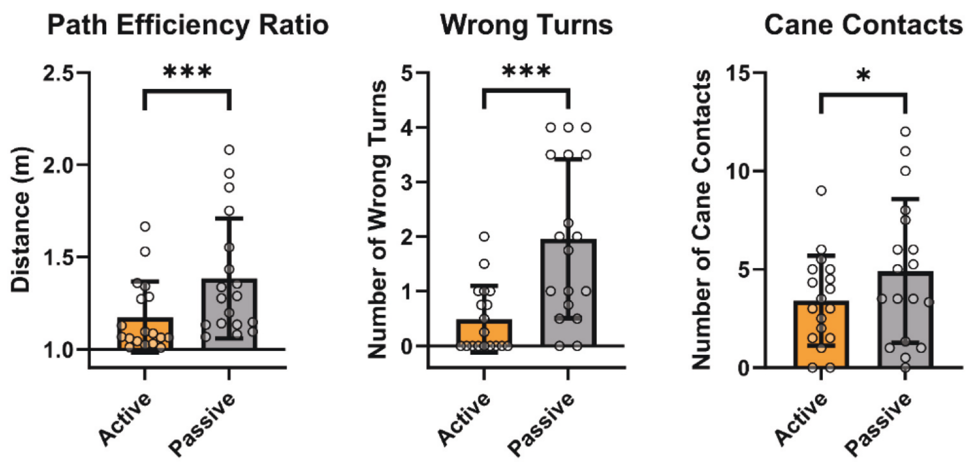

Figure 5. Spatial navigation performance of participants during active (UNav) or passive (SIPTD) conditions. For all spatial metrics, lower scores indicate better performance outcomes. From left to right, graphs show the path efficiency ratio (best pathing to actual pathing ratio), number of wrong turns taken by participants, and the number of times participants' canes contacted obstacles during the route. Dots indicate individual participant scores, error bars show ± 1SD, and significance ratings are as follows: * = $p < 0.05$, *** = $p < 0.001$.

Wrong turns

Participants made significantly fewer wrong turns in the active condition ($M = 0.49$) compared to the passive condition ($M = 1.96$), z = −3.52, $p < 0.001$, effect size = 0.88. This underscores the accuracy and reliability of the UNav system in guiding BLV users along the correct path.

Cane contacts

The number of cane contacts was significantly reduced in the active condition ($M = 3.40$) compared to the passive condition ($M = 4.91$), z = −2.43, $p = 0.015$, effect size = −0.61. This reduction may indicate that the UNav system enhances users' spatial awareness, reducing the need for physical contact with the cane for navigation assurance.

Temporal metrics

Total time taken

A significant difference was observed in the total time taken for navigation between the two conditions (see Figure 6). Participants navigating under the active condition took notably less time ($M = 155.0$ seconds) compared to those under the passive condition ($M = 185.1$ seconds), z = −3.22, $p = 0.001$, effect size = −0.76. The data suggest that the UNav system allows for faster navigation compared to traditional methods.

Idle time

There was a significant reduction in idle time for participants in the active condition ($M = 6.4$ seconds) compared to those in the passive condition ($M = 19.9$ seconds), z = −3.10, $p = 0.002$, effect size = −0.75. This reduction in idle time contributes to the overall efficiency and speed of navigation observed with the UNav system.

Number of idles

The active condition resulted in significantly fewer idle periods ($M = 1.26$) compared to the passive condition ($M = 2.63$), z = −2.72, $p = 0.007$, effect size = −0.66. This indicates that participants using the UNav system experienced fewer interruptions in their movement, resulting in a smoother navigation experience.

Gait speed

No significant effect was observed on the gait speed between the two conditions, with the active condition recording a mean gait speed of 0.75 m/s and the passive condition recording 0.80 m/s, z = −1.55, $p = 0.121$, effect size = −0.36.

User confidence

Help requests

Participants in the active condition sought help significantly fewer times ($M = 0.80$) than those in the passive condition ($M = 2.08$), z = −2.98, $p = 0.003$, effect size = −0.80 (see Figure 6). This suggests that the UNav system provides BLV users with increased independence and confidence in navigation.

Order effects

Condition ordering for total time, path efficiency, idle time

We evaluated whether initially learning a route with either UNav or SIPTD had differential effects on subsequent journeys with other conditions. For this, we compared passive-active vs active-passive trials for total time taken, users' path efficiency scores, and their total idle time (see Figure 7). These were chosen as they were continuous variables which previously had significant differences between UNav and SIPTD. We found that only the total time taken had a significant ordering effect, $t(10) = 2.324$, $p = 0.043$, $d = -0.701$, with a bigger change in performance from passive-active ($M = 40.00$) than active-passive ($M = -2.73$). This indicates that subjects using SIPTD were only as fast as UNav, *after* they had previously used UNav for that route.

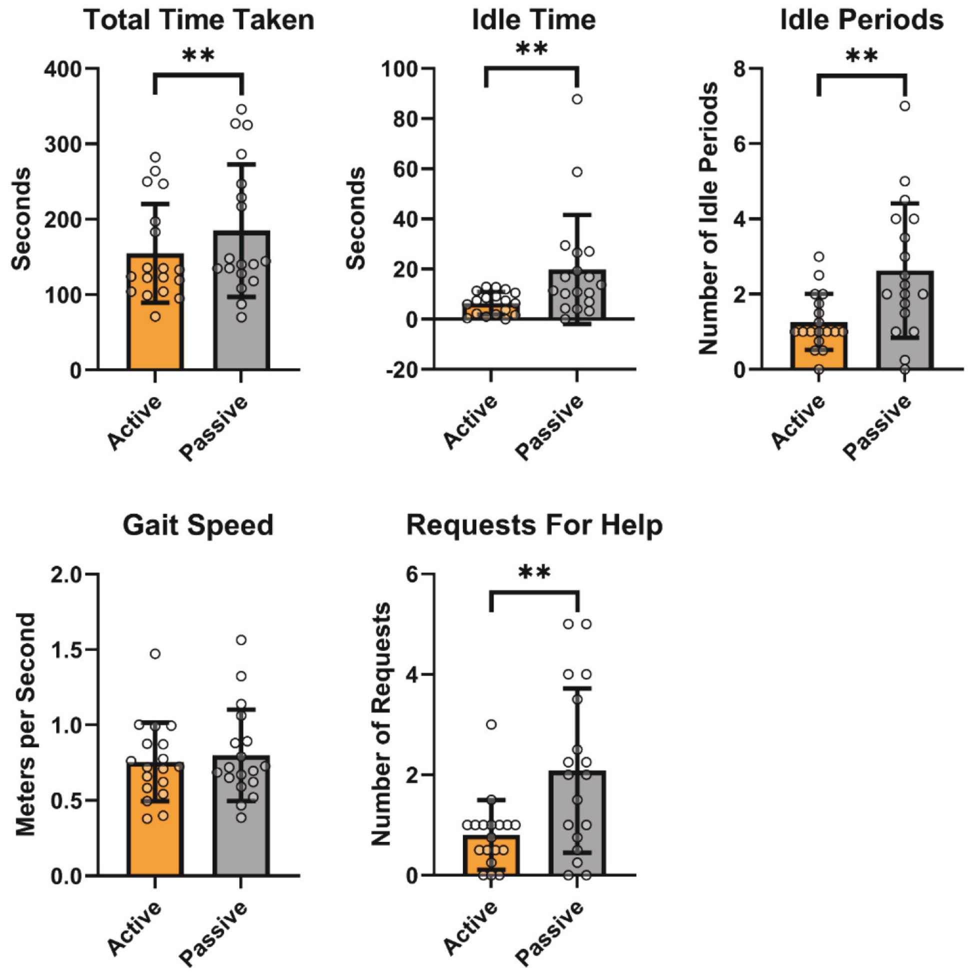

Figure 6. Temporal performance and user confidence metrics during navigation within active (UNav) and passive (SIPTD) conditions. For all temporal and user confidence metrics, lower scores indicate better performance outcomes, with the exception of gait speed. From left to right along the top row, graphs show the mean total time taken by participants to complete the navigation routes, the mean amount of time spent idling during routes, and the mean number of times participants stopped to idle. Along the bottom row, from left to right, graphs show participant's mean gait speed, and the mean number of requests for help from the chaperone per route. Dots indicate individual participant scores, error bars show ± 1SD, and significance ratings are as follows: * = $p < 0.05$, ** = $p < 0.01$.

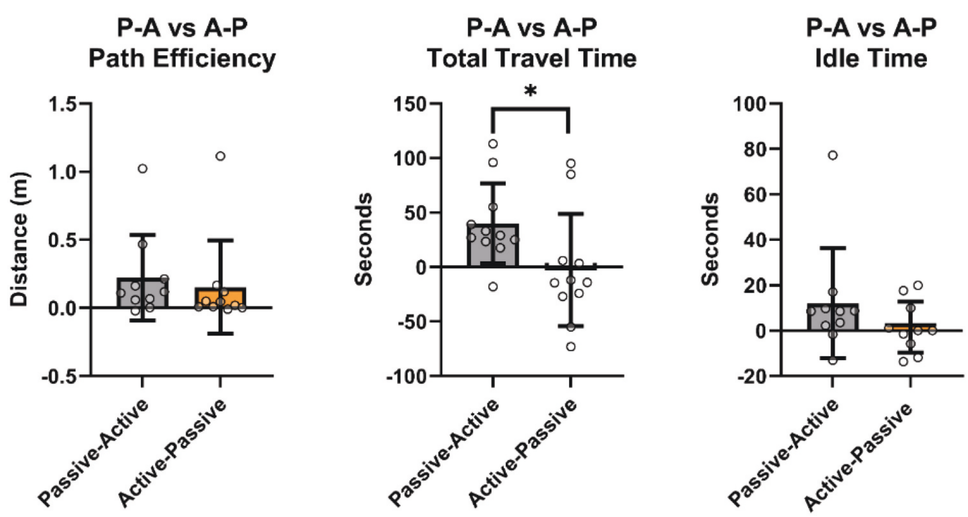

Figure 7. Comparison of order effects for passive-active vs active-passive ordering. For all metrics, lower scores indicate better performance. Each graph shows the mean difference between the initial condition followed by the subsequent condition. Values above zero indicate that the subsequent condition had better performance, values below zero indicate the initial condition had better performance for the same navigation route. From left to right, graphs show the ordering effects for the mean path efficiency, mean travel time, and mean idle time per route. Dots indicate individual participant scores, error bars show ± 1SD, and significance ratings are as follows: * = $p < 0.05$.

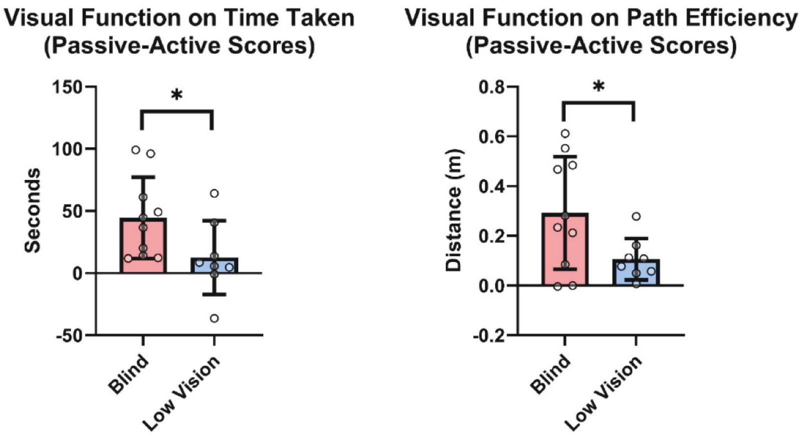

Figure 8. Comparison of subject's visual function (VF) scores on differences between passive and active conditions for total time taken (left) and path efficiency (right) metrics. Subjects are grouped into either "blind" (VF = 6) or "low vision" (VF ≤ 5). Scores above zero indicate that subjects' passive scores were higher (indicating worse performance) than their active scores. The higher the score, the more of a difference there was between their passive and active trials. Results indicate that the blind group had a significantly larger difference between passive and active trials for total time taken and path efficiency, indicating that they benefitted more from UNav in the active condition than the low vision group. Dots indicate individual participant scores, error bars show ± 1SD, and significance ratings are as follows: * = $p < 0.05$.

Effect of visual function

We evaluated whether subjects' visual function scores affected the degree to which subjects benefitted from the use of UNav over SIPTD. Here, the subjects were grouped into "blind" (VF = 6, $n = 10$) or "low vision" (VF ≤ 5, $n = 8$) and compared using different scores (passive-active scores) for all metrics except navigation success. We found a significant effect of the group on the total time taken and path efficiency (see Figure 8).

Total time

For the total time taken, the blind group had a significantly larger difference between passive and active trials ($M = 44.4$ seconds) than the low vision group (M = 12.4 seconds), $t(16) = 2.150$, $p = 0.047$, $d = 1.027$. This indicates that subjects who were blind benefited more from UNav assistance in terms of reducing their total travel time than those who had some remaining visual function.

Path efficiency

For the subject's path efficiency score, the blind group had a significantly larger difference between passive and active trials ($M = 0.292$) than the low vision group ($M = 0.106$), $t(16) = 2.191$, $p = 0.044$, $d = 1.197$. This indicates that subjects who were blind benefited more from UNav assistance in terms of improving their pathing efficiency than those who had some remaining visual function.

Discussion

The results of this study provide compelling insights into the advantages of the UNav wayfinding system in ecologically valid test environments over traditional in-person instructional guidance for PBLV. *Our data suggest that the UNav system offers a wide range of spatial, temporal and confidence benefits to the user, which should significantly enhance their navigation experience*. Here we will discuss how UNav likely contributes to each domain, spatial, temporal, and confidence, in turn. We will also explore how UNav's feedback could assist in constructing accurate mental maps of the environment and compare UNav with other EOAs in terms of what their specific approaches require to support navigation and their practical implications with explicit attention paid to cost, scale, and usability. Finally, we will discuss broader impacts of navigation tools for societal access, education, and employment opportunities.

Spatial performance and user confidence

We found that in terms of spatial performance, *UNav users reached their destination using more efficient pathing (resulting in shorter distances traveled), made fewer wrong turns, and required fewer cane contacts with walls and/or obstacles*. This indicates that participant's wayfinding was overall more efficient using the UNav system relative to chaperone instructions. Several factors may contribute to this, including UNav providing step-by-step route instructions as well as the consistency in the timing and style of spatial instructions provided. Chaperones providing the whole set of route instructions only once may have been more mentally taxing for working memory and leave room for subjective uncertainty, leading to more requests for help. That said, we balanced this known limitation with short trips with minimal complexity and a cap on total turns. In contrast, in active conditions, users may be more confident when trip support is delivered in more comprehensible steps, segment by segment. In real-world scenarios, it should also be considered that chaperones may not always be available (Dove et al., 2022) and also be more inconsistent in their instructions, describing distances, landmarks and turns in various levels of detail (Brügger et al., 2019). In-person, chaperone-based instructions were standardized in this instance to mitigate these factors. While overall navigation pathing success rates were high for both UNav (100%) and SIPTD (97.5%), the similarities may be due to the limited length and complexity of the routes (2–4 turns only) alongside the ability to request help. These success rates may deviate further from longer, more complex routes, in which in-person help may be limited or not available. All considered, UNav's consistency and reliability in delivering optimal pathing for the participant may underlie how the participant is able to use this navigation tool more effectively and with more confidence.

Temporal performance

In terms of our temporal metrics, *subjects in the UNav condition took less time to reach their destination, idled fewer times, and idled for less time overall*. All these advantages were seen despite gait speed remaining the same across passive and active conditions. Some of these gains are likely related to better spatial efficiency and user confidence, as better pathing (given a consistent gait speed) should reduce overall time as well. However, this can only occur if UNav's feedback is presented in a timely manner for subjects to allow them to reach their preferred walking speeds rather than having to wait in place for the next set of instructions. Future EOA evaluations should consider these holistic spatial-temporal-confidence frameworks to evaluate EOAs (Kappers et al., 2022; Tao & Ganz, 2019). Subject demographics and abilities should also be considered when designing feedbacks, such as considering each individual's preferred walking speed, step length, preferred routes, ability to memorize each route segment, alongside other demographic details that correlate with visual disabilities, such as age or medical comorbidities that may also impair device interaction or mobility (Nair et al., 2022).

Spatial cognition with mental maps – Evidence from ordering effects

As the participant navigates unfamiliar spaces using instructions from UNav or a chaperone, the user is building up a mental map of the environment and improving their spatial cognition (Dodds et al., 1982; Giudice, 2018; Hersh, 2020). Relative to the rote memorization required by SIPTD, UNav's live-generated step-by-step instructions might be more beneficial in mental map creation – which could benefit future journeys. One way to test this is through examining ordering effects, in which the effect of traveling a path first with UNav might influence subsequent travel with SIPTD (or vice versa). We examined the ordering-effects on three characteristics – path efficiency, total travel time and total idle time (see Figure 7). We found that "total travel time" with SIPTD could be reduced and reach equivalent times to UNav guidance and that this occurred in our experiment's active-passive condition where participants initially learned the route with UNav. By contrast, total travel time with UNav is not affected by whether the route has been traveled before with SIPTD. This suggests that *route learning and mental mapping in the initial journey with UNav may have resulted in greater benefits to spatial cognition that manifested in future trials, even without UNav*. However, further studies are needed to disentangle the roles of prior route experience and UNav on this effect, which we explore further in the limitations section.

Outdoor and indoor EOAs – Feature comparison

UNav's unique feature set addresses core navigation issues identified as barriers to equitable societal participation. UNav's vision-based localization method also provides orientation information, allowing the user to know the cardinal direction they are facing, as well as the cardinal direction of the destination. This information is similar to the orientation and distance information provided as 3D sound cues in Microsoft's Soundscape App. However, since Soundscape uses GPS information, its orientation accuracy is variable and cannot be applied to indoor environments unlike UNav. Another EOA, GoodMaps, has two separate applications with one for outdoor navigation using GPS and another for indoor navigation using previously generated 3D maps. *Since UNav utilizes the same visual localization method for both outdoors and indoors, there is a seamless handoff when users transition between outdoors and indoors*. Interviews with participants indicate that seamless handoffs between locations or tasks are highly desired.

Another core feature of UNav is its independence from the requisite infrastructure in the environment. This is similar to our Commute Booster application that utilizes preexisting signage to help navigate users within subway stations [58]. However, many alternative indoor approaches require Bluetooth beacons or other infrastructure to be placed in the environment to localize the user or provide content (e.g. Right Hear, NaviLens – for a review see (Plikynas et al., 2020)). These infrastructure requirements may be costly, require planning permission, compete for wall space, and require maintenance. VPR techniques do not require any of these, and the visual features required to support routes within indoor environments are captured in the same amount of time it takes to walk through an environment as if on a tour. Alternative techniques such as LiDAR mapping, augmented reality, magnetometer readings, and additional physical infrastructure-based methods also require walking through the environment, but in a very specific and choreographed manner. This can take considerably longer and may suffer from mapping inconsistencies that require repeated data acquisition, i.e. re-mapping. This means thatUNavis easily and rapidly scalable for larger environments without costly or space-consuming physical infrastructure or without ambitious, choreographed mapping routines that require trained personnel.

Server-based (remote) vs device-based (local) processing

EOAs for indoor navigation not only vary in whether or not they require additional physical infrastructure in the environment (Nair et al., 2022) but also whether network connectivity is required to facilitate travel guidance (GoodMaps, 2024). The iteration of UNav used in the present study offloads client image processing, visual feature matching for user localization and orientation, topometric map hosting, and wayfinding calculations to the server. This approach can yield dividends in terms of accessibility to this service from a wider range of devices as well as lowering the computational, memory, and battery demands on the client's system (Azzino et al., 2023; Yuan et al., 2022). All said, this approach is dependent on network connectivity and server availability. However, client–server interactions can be optimized so that reductions in network connection quality have minimal impact on server-based AI performance (Azzino et al., 2023). Ensuring robust navigation assistance for when network connections or servers are limited or intermittently unavailable. While UNav's server-side processing is lightweight enough to run on a laptop (Dell m17 R3 with RTX 2080 Super GPU), we are making several further optimizations to make UNav as efficient as possible to run locally on smartphones or the VIS^{4}ION platform. These efficiency improvements include both support for parallel processing on multiple CPU/GPU threads and the dynamic loading of smaller "sub-maps" into system memory based around the users' current location and next steps (rather than full-floor maps). These optimizations will work to improve processing speeds as well as reduce both system memory requirements and narrow the visual feature matching search to further speed up user localization, orientation, and wayfinding guidance.

Multi-purpose assistive technologies – Blending ETAs & EOAs

UNav provides consistent, reliable, step-by-step instructions to support navigation. Since subjects can rely on this information, *UNav users can go at their own pace and not have to memorize extensive lists of instructions and landmarks, reducing users' cognitive load*. As such, UNav users should be able to perceive and learn additional information about the environment to improve spatial cognition. To better facilitate mental map building, it may prove beneficial for future versions of UNav's verbal feedback to incorporate BLV's unique priorities within the environment. This includes adding specific sensory cues (e.g. floor texture), user notes or calling out additional landmarks in a particular rank order. As observed earlier, effective mental map (Hersh, 2020) building could create benefits for the user that extend beyond their direct use of the UNav system. UNav's feedback could be supplemented with hazard notifications of objects in the users' path (Hao et al., 2023), cross-walk notifications (Li et al., 2020), guidance to door handles (Gui et al., 2019), additional context from vision-based language models (e.g. Hao et al., 2024), alternative modalities such as via haptic feedback (Boldini et al., 2023), or provide trip planning guidance in virtual reality prior to real-world traveling (Ricci, Boldini, Beheshti, et al., 2023; Ricci, Boldini, Ma, et al., 2023; Ricci, Boldini, Rizzo, et al., 2023). This may facilitate user safety, ability, confidence, as well as improved spatial intelligence and situational awareness.

Limitations

In the present study, subjects completed routes twice, once with UNav and once with chaperone instructions. While this provided insights on how initially learning a route with one approach influenced learning with another (see ordering effects), the overall route learning effects may have reduced the observed differences between these approaches. This can be addressed in future studies with a careful selection of equivalent routes that are equally unfamiliar to subjects but have the exact same distance and turns required but are only used once by each approach.

We previously noted that route completion time by chaperone instructions (SIPTD) was only as fast as UNav *after* subjects initially learned the route using UNav. However, it is difficult to determine to what extent this was due to having completed the route before or is due to UNav's approach to presenting information. If future studies contain a passive–passive condition, then contrasting its results with the active-passive condition would help clarify the exact benefits of initially learning a route with UNav on subsequent journeys without UNav.

We interpret our ordering-effect results as evidence of UNav improving the subject's spatial cognition of the environment, which facilitates subsequent movement through the space. However, further studies could explore this interpretation through having subjects convey their understanding of the space after navigating it with either method for the first time (Giudice et al., 2011). Finally, in this study, we only tested short-range navigation (50–200 m, 2–4 turns); however, differences between UNav and chaperone instructions may differ further during longer or more complicated routes.

Spatial equity and societal access: A navigational priority

Spatial equity represents the goal of ensuring access to various spaces, communicating how they are distributed, and facilitating social inclusion for all activities possible within them (Cass et al., 2005). As such, EOAs can provide additional spatial equity for PBLV by providing more detailed spatial descriptions, detailed routing options, the location of nearby points of interest (e.g. shops, public bathrooms and transport hubs) as well as the various activities and goods available. This could result in EOAs describing the environment according to the users' preference, in addition to providing their selected destination. Accessibility to navigation is also an essential aspect of ensuring societal access for PBLV. The added difficulties for PBLV in commuting to and from work, effectively navigating inside the workplace, and initial worries that recruiters may have about workplace accommodations all play a role in employment barriers (Crudden & McBroom, 1999; U.S. Bureau of Labor Statistics, 2023). In the US, the unemployment rate for PBLV stands at ~60% relative to 4% for persons without disabilities (PwoD). These inequities extend to accessing healthcare services (Spencer et al., 2009), universities (Croft, 2020), and social spaces (Robertson, 2023; Wolffe & Sacks, 1997) – resulting in worse health outcomes, lower bachelor-degree attainment (15% relative to 41% PwoD), lower earning power, and fewer face-to-face social interactions relative to sighted peers (Wolffe & Sacks, 1997). These are substantial factors that add to both the societal but also financial cost associated with blindness and low vision (Rein et al., 2022). Accessible navigation solutions for PBLV are likely to help mitigate these societal inequities through facilitating access to educational, employment, social, and additional recreational opportunities.

Conclusion

Overall, we demonstrate that the UNav wayfinding system shows a wide range of significant benefits to the end user in terms of spatial, temporal, and confidence metrics over standard in-person travel directions when navigating through unfamiliar environments. Furthermore, we found that learning a route with UNav had persistent benefits, as users improved their travel time from standard in-person travel directions only after the user had previously learnt the route with UNav. These findings with UNav underscore the importance of consistent feedback from computer-based virtual guidance in spatial cognition and mental map building. Furthermore, UNav also serves as a suitable platform technology to explore further optimizations of feedback for both navigation instructions and overall environmental understanding, supporting multi-purpose assistive technology for better spatial intelligence and situational awareness. The infrastructure-free feature set of UNav also provides unique answers to problems faced by many alternative technologies, making it more widely accessible, cost-effective, and scalable. As a result, UNav has the potential to provide meaningful advances in accessible navigation for persons who are blind or have low vision to increase equitable access to a wide range of indoor and outdoor spaces, improving employment, healthcare, and educational opportunities both domestically and abroad.

Disclosure statement

NYU, Anbang Yang, John-Ross Rizzo, and Chen Feng have financial interests in the intellectual property described. We are affiliated with New York University, which is the applicant of the patent. NYU, as well as the inventors, hold equity and advisory positions in the related entity, potentially benefiting from the patent's commercialization. NYU has licensed the patent to Tactile Navigation Tools, where NYU, John-Ross Rizzo, and Todd E. Hudson are equity holders and advisors of said company.

Funding

The work was supported by the National Science Foundation [ITE2345139, DUE2129076, CNS1952180]; National Institutes of Health [R33EY033689]. The content is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health and the National Science Foundation. This work was also partly supported by Mahidol University and the Office of the Ministry of Higher Education, Science, Research, and Innovation under the Reinventing University project: the Center of Excellence in AI-Based Medical Diagnosis (AI-MD) sub-project.

References

Ahmetovic, D., Gleason, C., Ruan, C., Kitani, K., Takagi, H., & Asakawa, C. (2016, September). NavCog: A navigational cognitive assistant for the blind. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, in MobileHCI '16 (pp. 90–99). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2935334.2935361

Azzino, T., Mezzavilla, M., Rangan, S., Wang, Y., & Rizzo, J.-R. (2023, November 23). 5G edge vision: Wearable assistive technology for people with blindness and Low vision. *arXiv*. https://doi.org/10.48550/arXiv.2311.13939

Ball, E. M. (2008). Electronic travel aids: An assessment. In M. A. Hersh & M. A. Johnson (Eds.), *Assistive technology for visually impaired and blind people* (pp. 289–321). Springer. https://doi.org/10.1007/978-1-84628-867-8_9

Bohonos, S., Lee, A., Malik, A., Thai, C., & Manduchi, R. (2007, June). Universal real-time navigational assistance (URNA): An urban bluetooth beacon for the blind. *Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments*, in HealthNet '07 (pp. 83–88). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/1248054.1248080

Boldini, A., Rizzo, J.-R., & Porfiri, M. (2023, September). Macro-fiber composite-based tactors for haptic applications. *IEEE Transactions on Haptics*, *16*(3), 436–448. https://doi.org/10.1109/TOH.2023.3308789

Bourne, R. R. A., Sukudom, P., Foster, P.J., Tantisevi, V., Jitapunkul, S., Lee, P.S., Johnson, G.J., & Rojanapongpun, P. (2003, September). Prevalence of glaucoma in Thailand: A population based survey in Rom Klao District, Bangkok. *British Journal of Ophthalmology*, *87*(9), 1069–1074. https://doi.org/10.1136/bjo.87.9.1069

Brügger, A., Richter, K.-F., & Fabrikant, S. I. (2019, February). How does navigation system behavior influence human behavior? *Cognitive Research: Principles and Implications*, *4*(1), 5. https://doi.org/10.1186/s41235-019-0156-5

Cass, N., Shove, E., & Urry, J. (2005, August). Social exclusion, mobility and access. *The Sociological Review*, *53*(3), 539–555. https://doi.org/10.1111/j.1467-954X.2005.00565.x

Chen, C., Liu, X., Li, Y., Ding, L., & Feng, C. (2023). DeepMapping2: Self-supervised large-scale LiDAR map optimization. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada (pp. 9306–9316).

Cheraghi, S. A., Namboodiri, V., & Walker, L. (2017, March). GuideBeacon: Beacon-based indoor wayfinding for the blind, visually impaired, and disoriented. *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)* (pp. 121–130). https://doi.org/10.1109/PERCOM.2017.7917858

Croft, E. (2020, December). Experiences of visually impaired and blind students in UK higher education: An exploration of access and participation. *Scandinavian Journal of Disability Research*, *22*(1, Art. no. 1), 382–392. https://doi.org/10.16993/sjdr.721

Crudden, A., & McBroom, L. W. (1999, June). Barriers to employment: A survey of employed persons who are visually impaired. *Journal of Visual Impairment & Blindness*, *93*(6), 341–350. https://doi.org/10.1177/0145482X9909300602

Dodds, A. G., Howarth, C. I., & Carter, D. C. (1982, January). The mental maps of the blind: The role of previous visual experience. *Journal of Visual Impairment & Blindness*, *76*(1), 5–12. https://doi.org/10.1177/0145482X8207600102

dos Santos, A. D. P., Ferrari, A. L. M., Medola, F. O., & Sandnes, F. E. (2022, February). Aesthetics and the perceived stigma of assistive technology for visual impairment. *Disability and Rehabilitation Assistive Technology*, *17*(2), 152–158. https://doi.org/10.1080/17483107.2020.1768308

Dove, G., Fernando, A., Hertz, K., Kim, J., Rizzo, J.-R., Seiple, W. H., & Nov, O. (2022, November). Digital technologies in orientation and mobility instruction for people who are blind or have Low vision. *Proceedings of the ACM on Human-Computer Interaction*, *6* (CSCW2), :509:1–:509:25. https://doi.org/10.1145/3555622

Fallah, N., Apostolopoulos, I., Bekris, K., & Folmer, E. (2012, May). The user as a sensor: Navigating users with visual impairments in indoor spaces using tactile landmarks. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, in CHI '12 (pp. 425–432). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2207676.2207735

Feng, J., Beheshti, M., Philipson, M., Ramsaywack, Y., Porfiri, M., & Rizzo, J. R. (2023). Commute booster: A mobile application for first/last mile and middle mile navigation support for people with blindness and Low vision. *IEEE Journal of Translational Engineering in Health and Medicine*, *11*, 523–535. https://doi.org/10.1109/JTEHM.2023.3293450.

Gilbert, C., & Foster, A. (1993, August). Causes of blindness in children attending four schools for the blind in Thailand and the Philippines. *International Ophthalmology*, *17*(4), 229–234. https://doi.org/10.1007/BF01007745

Giudice, N. A. (2018). Navigating without vision: Principles of blind spatial cognition. In D. R. Montello (Ed.), *Handbook of behavioral and cognitive geography* (pp. 260–288). Edward Elgar Publishing.

Retrieved April 10, 2024, [Online]. https://www.elgaronline.com/edcollchap/edcoll/9781784717537/9781784717537.00024.xml

Giudice, N. A., Betty, M. R., & Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: Evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *37*(3), 621–634. https://doi.org/10.1037/a0022331

Giudice, N. A., Whalen, W. E., Riehle, T. H., Anderson, S. M., & Doore, S. A. (2019, March). Evaluation of an accessible, real-time, and infrastructure-free indoor navigation system by users who are blind in the Mall of America. *Journal of Visual Impairment & Blindness*, *113*(2), 140–155. https://doi.org/10.1177/0145482X19840918

GoodMaps. (2024, March, 30). *GoodMaps. Mapping the future with indoor navigation*. https://goodmaps.com/

Goswami, R. G., Amith, P.V., Hari, J., Dhaygude, A., Krishnamurthy, P., Rizzo, J., Tzes, A., & Khorrami, F. (2023, February). Efficient real-time localization in prior indoor maps using semantic SLAM. *2023 9th International Conference on Automation, Robotics and Applications (ICARA)* (pp. 299–303). https://doi.org/10.1109/ICARA56516.2023.10125919

Gui, W., Li, B., Yuan, S., Rizzo, J.R., Sharma, L., Feng, C., Tzes, A., & Fang, Y. (2019, November). An assistive low-vision platform that augments spatial cognition through proprioceptive guidance: Point-to-Tell-and-Touch. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 3817–3822). https://doi.org/10.1109/IROS40897.2019.8967647

Hakobyan, L., Lumsden, J., O'Sullivan, D., & Bartlett, H. (2013, November). Mobile assistive technologies for the visually impaired. *Survey of Ophthalmology*, *58*(6), 513–528. https://doi.org/10.1016/j.survophthal.2012.10.004

Han, Y., Beheshti, M., Jones, B., Hudson, T. E., Seiple, W. H., & Rizzo, J.-R. (2023, May). Wearables for persons with blindness and low vision: Form factor matters. *Assistive Technology*, *36*(1), 60–63. https://doi.org/10.1080/10400435.2023.2205490

Hao, Y., Feng, J., Rizzo, J.-R., Wang, Y., & Fang, Y. (2023). Detect and approach: Close-range navigation support for people with blindness and Low Vision. In L. Karlinsky, T. Michaeli, & K. Nishino (Eds.), *Lecture Notes in Computer Science. Computer vision – ECCV 2022 workshops* (pp. 607–622). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-25075-0_41

Hao, Y., Yang, F., Huang, H., Yuan, S., Rangan, S., Rizzo, J. R., Wang, Y., & Fang, Y. (2024). A multi-modal foundation model to assist people with blindness and low vision in environmental interaction. *Journal of Imaging*, *10*(5), 103. https://doi.org/10.3390/jimaging10050103

Hersh, M. (2020, April). Mental maps and the use of sensory information by blind and partially sighted people. *ACM Transactions on Accessible Computing*, *13*(2), :6:1–:6:32. https://doi.org/10.1145/3375279

Hersh, M. (2022, January). Wearable travel aids for blind and partially sighted people: A review with a focus on design issues. *Sensors*, *22*(14, Art. no. 14), 5454. https://doi.org/10.3390/s22145454

Isipradit, S., Sirimaharaj, M., Charukamnoetkanok, P., Thonginnetra, O., Wongsawad, W., Sathornsumetee, B., Somboonthanakij, S., Soomsawasdi, P., Jitawatanarat, U., Taweebanjongsin, W., Arayangkoon, E., Arame, P., Kobkoonthon, C., & Pangputhipong, P. (2014, December). The first rapid assessment of avoidable blindness (RAAB) in Thailand. *PLOS ONE*, *9*(12), e114245. https://doi.org/10.1371/journal.pone.0114245

Jain, M., & Patel, W. (2023, April). Review on LiDAR-based navigation systems for the visually impaired. *SN Computer Science*, *4*(4), 323. https://doi.org/10.1007/s42979-023-01735-y

Jeamwatthanachai, W., Wald, M., & Wills, G. (2019a, May). Indoor navigation by blind people: Behaviors and challenges in unfamiliar spaces and buildings. *The British Journal of Visual Impairment*, *37*(2), 140–153. https://doi.org/10.1177/0264619619833723

Jeamwatthanachai, W., Wald, M., & Wills, G. (2019b, January). Spatial representation framework for better indoor navigation by people with visual impairment. *Journal of Enabling Technologies*, *13*(4), 212–227. https://doi.org/10.1108/JET-12-2018-0068

Jenchitr, W., Ruamviboonsuk, P., Sanmee, A., & Pokawattana, N. (2011, February). Prevalence of age-related macular degeneration in Thailand. *Ophthalmic Epidemiology*, *18*(1), 48–52. https://doi.org/10.3109/09286586.2010.545502

Kandalan, R. N., & Namuduri, K. (2019, June 13). A comprehensive survey of navigation systems for the visual impaired. *arXiv*. https://doi.org/10.48550/arXiv.1906.05917

Kappers, A. M. L., Oen, M. F. S., Junggeburth, T. J. W., & Plaisier, M. A. (2022, December). Hand-held haptic navigation devices for actual walking. *IEEE Transactions on Haptics*, *15*(4), 655–666. https://doi.org/10.1109/TOH.2022.3209350

Kruk, M. E., & Pate, M. (2020, January). The lancet global health commission on high quality health systems 1 year on: Progress on a global imperative. *The Lancet Global Health*, *8*(1), e30–e32. https://doi.org/10.1016/S2214-109X(19)30485-1

Li, X., Cui, H., Rizzo, J.-R., Wong, E., & Fang, Y. (2020). Cross-safe: A computer vision-based approach to make all intersection-related pedestrian signals accessible for the visually impaired. In K. Arai & S. Kapoor (Eds.), *Advances in Intelligent Systems and Computing. Advances in computer vision* (pp. 132–146). Springer International Publishing. https://doi.org/10.1007/978-3-030-17798-0_13

Mahdi, E., & Xinming, H. (2022, August 05). A survey on visual map localization using LiDARs and cameras. *arXiv*. https://doi.org/10.48550/arXiv.2208.03376

McGrath, C., & Astell, A. (2017, February). The benefits and barriers to technology acquisition: Understanding the decision-making processes of older adults with age-related vision loss (ARVL). *The British Journal of Occupational Therapy*, *80*(2), 123–131. https://doi.org/10.1177/0308022616667959

Nair, V., Olmschenk, G., Seiple, W. H., & Zhu, Z. (2022, May). ASSIST: Evaluating the usability and performance of an indoor navigation assistant for blind and visually impaired people. *Assistive Technology*, *34*(3), 289–299. https://doi.org/10.1080/10400435.2020.1809553

National Research Council. (2002). *Visual impairments: Determining eligibility for social security benefits*. National Academies Press.

nPerf. (2024, April 10). *5G coverage map worldwide- nPerf.com*. https://www.nperf.com/en/map/5g

Okonji, P. E., & Ogwezzy, D. C. (2019). Awareness and barriers to adoption of assistive technologies among visually impaired people in Nigeria. *Assistive Technology*, *31*(4), 209–219. https://doi.org/10.1080/10400435.2017.1421594

Otero, R., Lagüela, S., Garrido, I., & Arias, P. (2020, December). Mobile indoor mapping technologies: A review. *Automation in Construction*, *120*, 103399. https://doi.org/10.1016/j.autcon.2020.103399

Pallant, Julie and Manual. (2011). *SPSS survival, "A step by step guide to data analysis using spss* (4th ed.). Aust. Allen Unwin.

Pan, L., Pollefeys, M., & Larsson, V. (2022). Camera pose estimation using implicit distortion models. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, Louisiana (pp. 12819–12828).

Pascolini, D., & Mariotti, S. P. (2012, May). Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, *96*(5), 614–618. https://doi.org/10.1136/bjophthalmol-2011-300539

Plikynas, D., Žvironas, A., Budrionis, A., & Gudauskis, M. (2020, January). Indoor navigation systems for visually impaired persons: Mapping the features of existing technologies to user needs. *Sensors*, *20*(3, Art. no. 3), 636. https://doi.org/10.3390/s20030636

Prabhasawat, P., Trethipwanit, K., Prakairungthong, N., Narenpitak, S., Jaruroteskulchai, S., & Anantachai, J. (2007). Causes of corneal blindness: A multi-center retrospective review. *Medical Journal of the Medical Association of Thailand*, *90*(12), 2651.

Ramalingam, B., Le, A. V., Lin, Z., Weng, Z., Mohan, R. E., & Pookkuttath, S. (2022, September). Optimal selective floor cleaning using deep learning algorithms and reconfigurable robot hTetro. *Scientific Reports*, *12*(1, Art. no. 1). https://doi.org/10.1038/s41598-022-19249-7

Rein, D. B., Wittenborn, J. S., Zhang, P., Sublett, F., Lamuda, P. A., Lundeen, E. A., & Saaddine, J. (2022, April). The economic burden of vision loss and blindness in the United States. *Ocular Diagnoses at*

Enrollment Ophthalmology, *129*(4), 369–378. https://doi.org/10.1016/j.ophtha.2021.09.010

Ricci, F. S., Boldini, A., Beheshti, M., Rizzo, J.-R., & Porfiri, M. (2023, June). A virtual reality platform to simulate orientation and mobility training for the visually impaired. *Virtual Reality*, *27*(2), 797–814. https://doi.org/10.1007/s10055-022-00691-x

Ricci, F. S., Boldini, A., Ma, X., Beheshti, M., Geruschat, D. R., Seiple, W. H., Rizzo, J.-R., & Porfiri, M. (2023). Virtual reality as a means to explore assistive technologies for the visually impaired. *PLOS Digital Health*, *2*(6), e0000275. https://doi.org/10.1371/journal.pdig.0000275

Ricci, F. S., Boldini, A., Rizzo, J.-R., & Porfiri, M. (2023, April). A multiplayer virtual reality platform to evaluate electronic travel aid performance for persons with blindness and low vision. In *Nano-, bio-, info-tech sensors, and wearable systems 2023* (pp. 11–17). SPIE. https://doi.org/10.1117/12.2658453

Rizzo, J.-R., Beheshti, M., Hudson, T. E., Mongkolwat, P., Riewpaiboon, W., Seiple, W., Ogedegbe, O. G., & Vedanthan, R. (2023, April). The global crisis of visual impairment: An emerging global health priority requiring urgent action. *Disability and Rehabilitation Assistive Technology*, *18*(3), 240–245. https://doi.org/10.1080/17483107.2020.1854876

Robertson, S. *Private: Chapter 22: Social interaction*. Retrieved November 12, 2023 [Online]. https://openpress.usask.ca/soc111/chapter/chapter-22-social-interaction/

Sarlin, P.-E., Cadena, C., Siegwart, R., & Dymczyk, M. (2019). From coarse to fine: Robust hierarchical localization at large scale. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, California (pp. 12716–12725).

Sato, D., Oh, U., Guerreiro, J., Ahmetovic, D., Naito, K., Takagi, H., Kitani, K. M., & Asakawa, C. (2019, August). NavCog3 in the wild: Large-scale blind indoor navigation assistant with semantic features. *ACM Transactions on Accessible Computing*, *12*(3), :14:1–:14:30. https://doi.org/10.1145/3340319

Senjam, S. S., Foster, A., Bascaran, C., & Vashist, P. (2019, October). Awareness, utilization and barriers in accessing assistive technology among young patients attending a low vision rehabilitation clinic of a tertiary eye care centre in Delhi. *Indian Journal of Ophthalmology*, *67* (10), 1548–1554. https://doi.org/10.4103/ijo.IJO_197_19

Sheng, D., Chai, Y., Li, X., Feng, C., Lin, J., Silva, C., & Rizzo, J.R. (2021, September). NYU-VPR: Long-term visual place recognition benchmark with view direction and data anonymization influences. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 9773–9779). https://doi.org/10.1109/IROS51168.2021.9636640

Spencer, C., Frick, K., Gower, E. W., Kempen, J. H., & Wolff, J. L. (2009, January). Disparities in access to medical care for individuals with vision impairment. *Ophthalmic Epidemiology*, *16*(5), 281–288. https://doi.org/10.1080/09286580902999439

Statista. (2024a, April 10). *Thailand: 5G average download speed of network providers 2024 | Statista*. https://www.statista.com/statistics/1242348/thailand-5g-average-download-speed-of-network-providers/

Statista. (2024b, April 10). *Thailand: Availability of 5G network by provider 2024 | Statista*. https://www.statista.com/statistics/1340972/thailand-availability-of-5g-network-by-provider/

Suebsin, C., & Gerdsri, N. (2012, July). The analysis of key factors affecting the success of technology adoption in healthcare organizations in Thailand. *2012 Proceedings of PICMET '12: Technology Management for Emerging Technologies* (pp. 1516–1526). Retrieved December 12, 2023 [Online]. https://ieeexplore.ieee.org/abstract/document/6304172?casa_token=Ib3QNcfuBicAAAAA:4l_a_vKBgslhMmwT9zJRSqVpNIoaAUXnCqIgsAfszBYjQ3MwK362orbCekMEVLhmG9igXERaCTIy

Tao, Y., & Ganz, A. (2019). Validation and optimization framework for indoor navigation systems using user comments in spatial-temporal context. *Institute of Electrical and Electronics Engineers Access*, *7*, 159479–159494. https://doi.org/10.1109/ACCESS.2019.2951014

U.S. Bureau of Labor Statistics. Retrieved November 26, 2023 [Online]. https://www.bls.gov/

Wolffe, K., & Sacks, S. Z. (1997, May). The lifestyles of blind, low vision, and sighted youths: A quantitative comparison. *Journal of Visual Impairment & Blindness*, *91*(3), 245–257. https://doi.org/10.1177/0145482X9709100310

Xu, P., Kennedy, G. A., Zhao, F.-Y., Zhang, W.-J., & Van Schyndel, R. (2023). Wearable obstacle avoidance electronic travel aids for blind and visually impaired individuals: A systematic review. *Institute of Electrical and Electronics Engineers Access*, *11*, 66587–66613. https://doi.org/10.1109/ACCESS.2023.3285396

Yang, A., Beheshti, M., Hudson, T. E., Vedanthan, R., Riewpaiboon, W., Mongkolwat, P., Feng, C., & Rizzo, J.-R. (2022, January). UNav: An infrastructure-independent vision-based navigation system for people with blindness and Low vision. *Sensors*, *22*(22, Art. no. 22), 8894. https://doi.org/10.3390/s22228894

Yang, A., Wang, Y., Rizzo, J.-R., & Feng, C. (2023, October 10). Distillation improves visual place recognition for low-quality queries. *arXiv*. https://doi.org/10.48550/arXiv.2310.06906

Yoon, C., Louie, R., Ryan, J., Vu, M., Bang, H., Derksen, W., & Ruvolo, P., et al. (2019, October). Leveraging augmented reality to create apps for people with visual disabilities: A case study in indoor navigation. *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, in ASSETS '19 (pp. 210–221). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3308561.3353788

Yuan, Z., Azzino, T., Hao, Y., Lyu, Y., Pei, H., Boldini, A., Mezzavilla, M., Beheshti, M., Porfiri, M., Hudson, T. E., Seiple, W., Fang, Y., Rangan, S., Wang, Y., & Rizzo, J.-R. (2022). Network-aware 5G edge computing for object detection: Augmenting wearables to "see" more, farther and faster. *Institute of Electrical and Electronics Engineers Access*, *10*, 29612–29632. https://doi.org/10.1109/ACCESS.2022.3157876