StreetNav: Leveraging Street Cameras to Support Precise Outdoor Navigation for Blind Pedestrians

Gaurav Jain
Columbia University
New York, NY, USA

Basel Hindi
Columbia University
New York, NY, USA

Zihao Zhang
Columbia University
New York, NY, USA

Koushik Srinivasula
Columbia University
New York, NY, USA

Mingyu Xie
Columbia University
New York, NY, USA

Mahshid Ghasemi
Columbia University
New York, NY, USA

Daniel Weiner[*]
Lehman College
New York, NY, USA

Sophie Ana Paris[*]
New York University
New York, NY, USA

Xin Yi Therese Xu[*]
Pomona College
Claremont, CA, USA

Michael Malcolm[*]
City Tech CUNY
New York, NY, USA

Mehmet Turkcan
Columbia University
New York, NY, USA

Javad Ghaderi
Columbia University
New York, NY, USA

Zoran Kostic
Columbia University
New York, NY, USA

Gil Zussman
Columbia University
New York, NY, USA

Brian A. Smith
Columbia University
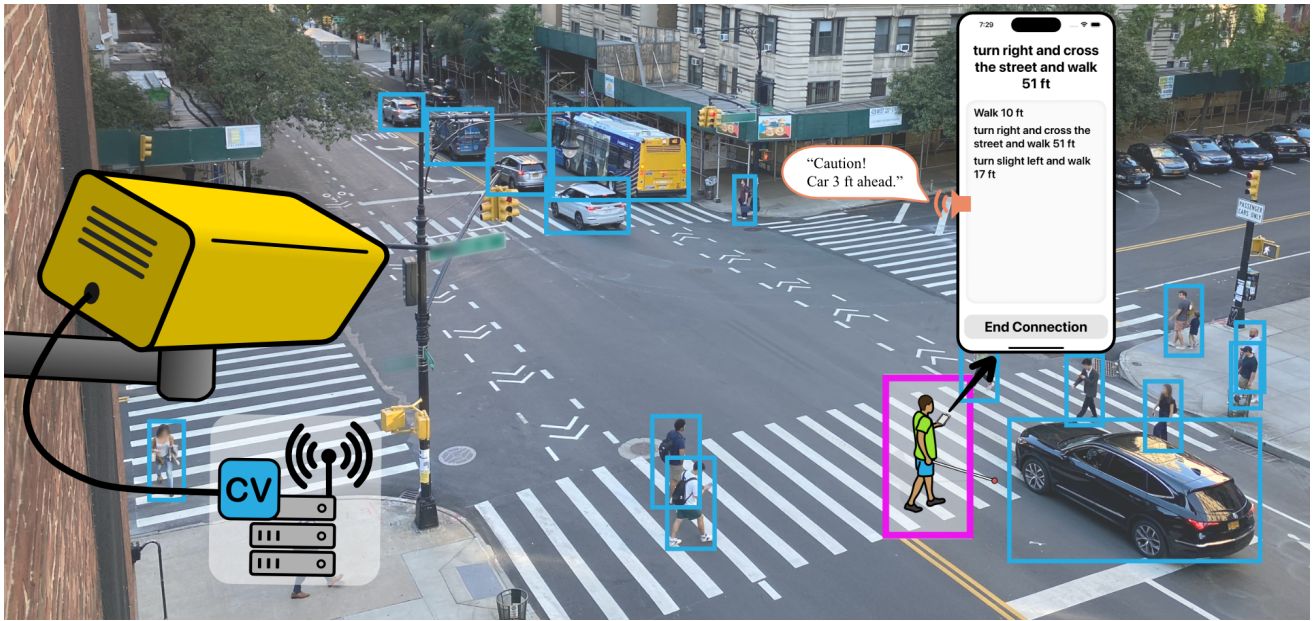New York, NY, USA

Figure 1: StreetNav is a system that explores the concept of repurposing existing street cameras to support precise outdoor navigation for blind and low-vision (BLV) pedestrians. It comprises two components: (i) a computer vision (CV) pipeline, and (ii) a companion smartphone app. The computer vision pipeline processes the street camera's video feeds and delivers real-time navigation feedback via the app. StreetNav offers precise turn-by-turn directions to destinations while also providing real-time, scene-aware assistance to prevent users from veering off course, alert them of nearby obstacles, and facilitate safe street crossings.

[*]This work was done while Daniel Weiner, Sophie Ana Paris, Xin Yi Therese Xu, and Michael Malcolm were interns at Columbia University.

ABSTRACT

Blind and low-vision (BLV) people rely on GPS-based systems for outdoor navigation. GPS's inaccuracy, however, causes them to veer off track, run into unexpected obstacles, and struggle to reach

precise destinations. While prior work has made precise navigation possible indoors via additional hardware installations, enabling precise navigation outdoors remains a challenge. Ironically, many outdoor environments of interest such as downtown districts are already instrumented with hardware such as street cameras. In this work, we explore the idea of repurposing street cameras for outdoor navigation, and investigate the effectiveness of such an approach. Our resulting system, StreetNav, processes the cameras' video feeds using computer vision and gives BLV pedestrians real-time navigation assistance. Our user evaluations in the COSMOS testbed with eight BLV pedestrians show that StreetNav guides them more precisely than GPS, but its performance is sensitive to lighting conditions and environmental occlusions. We discuss future implications for deploying such systems at scale.

CCS CONCEPTS

• **Human-centered computing → Accessibility systems and tools**.

KEYWORDS

Visual impairments, outdoor navigation, street camera, computer vision

1 INTRODUCTION

Outdoor navigation in unfamiliar environments is a major challenge for blind and low-vision (BLV) people. Among the many navigation systems that have been developed to assist BLV people outdoors, GPS-based systems are the most popular [28, 31, 41, 58, 64]. These systems, such as BlindSquare [41] and Microsoft Soundscape [28], guide users to a destination and notify them of surrounding points of interest (POIs). Despite GPS's undeniable impact in making outdoor environments navigable, its imprecision is a major limitation [56]. GPS precision can range from 5 meters at best to over tens of meters in urban areas with buildings and trees [23, 42, 65]. This imprecision causes BLV people to veer off track [49], run into unexpected obstacles [8, 50, 52], and struggle to reach precise destinations [56] when navigating outdoors.

Prior work on indoor navigation, on the contrary, has made precise navigation assistance possible for BLV people [2, 19, 34, 44, 57]. Most approaches do so by installing a dense network of additional hardware, such as Bluetooth [2] or WiFi [19] beacons, to precisely locate a user's position. Retrofitting outdoor environments with additional hardware, however, is not feasible due to the vast scale and complex nature of outdoor spaces. It would require extensive financial investments and coordination with city authorities to install and maintain such specialized hardware, which may not be possible.

Ironically, many outdoor environments of interest, such as urban districts and downtown areas, are already instrumented with hardware that has the potential to help, including street cameras, traffic sensors, and other urban infrastructure components. Street cameras, in particular, are increasingly being installed in cities for public safety, surveillance, and traffic management-related applications [4, 11, 18, 36, 39]. Although these pre-existing street cameras have been deployed for purposes unrelated to accessibility, their potential for facilitating navigation assistance for BLV people remains largely untapped.

In this work, we explore the idea of leveraging existing street cameras to support outdoor navigation assistance, and we investigate the effectiveness of such an approach. We seek to answer the following research questions:

RQ1. What challenges do BLV people face when navigating outdoors using GPS-based systems?

RQ2. How should street camera-based systems be designed to address BLV people's challenges in outdoor navigation?

RQ3. To what extent do street camera-based systems address BLV people's challenges in outdoor navigation?

To answer RQ1, we conducted formative interviews with six BLV pedestrians and discovered the challenges BLV people face when navigating outdoors using GPS-based systems. Our participants reported challenges in following GPS's routing instructions through complex environment layouts, avoiding unexpected obstacles while using assistive technology, and crossing streets safely.

To answer RQ2, we developed *StreetNav*, a system that leverages a street camera to support precise outdoor navigation for BLV pedestrians. As Figure 1 illustrates, StreetNav comprises two key components: (i) *a computer vision pipeline* and (ii) *a companion smartphone app*. The computer vision pipeline processes the street camera's video feed and delivers real-time navigation assistance to BLV pedestrians via the smartphone app. StreetNav offers precise turn-by-turn directions to destinations while also providing real-time, scene-aware assistance to prevent users from veering off course, alert them of nearby obstacles, and facilitate safe street crossings. We developed StreetNav using the NSF PAWR COSMOS wireless edge-cloud testbed [55, 68]. StreetNav uses one of COSMOS testbed's street cameras mounted on the second floor of Columbia University's Mudd building in New York City (NYC), which faces a four-way street intersection.

To answer RQ3, we conducted user evaluations involving eight BLV pedestrians who navigated routes with both StreetNav and BlindSquare [41], a popular GPS-based navigation app especially designed for BLV people. Our findings reveal that StreetNav offers significantly greater precision in guiding pedestrians compared to BlindSquare. Specifically, StreetNav guided participants to within an average of 2.9 times closer to their destination and reduced veering off course by over 53% when compared to BlindSquare. This substantial improvement was reflected in the unanimous preference of all participants for StreetNav over BlindSquare in a forced ranking. Our evaluation, however, also revealed technical considerations related to StreetNav's performance, notably its sensitivity to lighting conditions and environmental occlusions. We discuss the future implications of our findings in the context of deploying street camera-based systems at scale for outdoor navigation assistance.

In summary, we contribute (1) a formative study of BLV people's challenges in outdoor navigation using GPS-based systems, (2) the StreetNav system through which we explore the concept of repurposing street cameras for precise outdoor navigation assistance, and (3) a user evaluation of StreetNav.

2 RELATED WORK

Our work builds on the following three main research threads: (i) outdoor navigation approaches, (ii) overhead camera-based robot navigation, and (iii) indoor navigation approaches.

Outdoor Navigation Approaches. Existing approaches for outdoor navigation primarily rely on GPS-based navigation systems for guiding users to the destination and providing information about nearby POIs [28, 31, 41, 58, 64]. BlindSquare[41], for instance, utilizes the smartphone's GPS signal to determine the user's location and then provides the direction and distance to the destination, gathered from Foursquare and Open Street Map. The GPS signal, however, offers poor precision with localization errors as big as tens of meters [2, 23, 42, 69]. The accuracy is lower in densely populated cities [66], which is even more concerning given that a disproportionately high percentage of BLV people live in cities [27]. Despite GPS-based systems' undeniable impact on helping BLV people in outdoor navigation, their low precision and inability to provide real-time support for avoiding obstacles and veering off the path limits their usability as a standalone navigation solution. Our preliminary work on StreetNav [29] introduced the alternative of leveraging street cameras for outdoor navigation assistance. In this work, we investigate street cameras' potential for providing precise and real-time navigation assistance by performing a user experience evaluation of StreetNav.

Another approach for outdoor navigation has explored developing personalized, purpose-built, assistive devices that support BLV people with scene-aware aspects of outdoor navigation, such as crossing streets [26, 37, 61], recording routes [69], and avoiding obstacles [15, 16, 32, 38, 54, 67]. While these solutions address some of the precise and real-time aspects of BLV people's outdoor navigation, support for point-to-point navigation is missing. Consequently, they do not offer a comprehensive, all-in-one solution for outdoor navigation. Furthermore, these systems place the burden of purchasing costly devices onto the BLV users. Our work, by contrast, explores the possibility of using existing street cameras to provide a comprehensive solution for outdoor navigation. We investigate repurposing existing hardware in outdoor environments to support accessibility applications, thus imbuing accessibility within the city infrastructure directly, and adding no additional cost to the BLV user.

Overhead Camera-based Robot Navigation. A parallel research space to street cameras for blind navigation is robot navigation using overhead cameras. One common subspace within this field is sensor fusion for improved mapping. Research in this space focuses on fusing information between sighted "guide" robots and overhead cameras [10], fusing multiple camera views for improved tracking [10, 48, 51], and improving homography for robust mapping, independent of camera viewing angle [59, 60]. Another challenge tackled within this space is robot path planning. Research in this

space aims to improve path planning algorithms [10, 48, 60], assign navigational tasks to robot assistants [10, 48], and address the balance between obstacle avoidance and path following [10, 60]. While prior work on robot navigation using fixed cameras explores the research space of automating "blind" robot navigation, our work explores how fixed cameras, specifically street cameras, could be repurposed to support navigation for blind pedestrians. Our work considers BLV users' needs and preferences around outdoor navigation to design and develop a system that can offer precise navigation assistance.

Indoor Navigation Approaches. Prior work in indoor navigation assistance has made significant progress through the utilization of various localization technologies, which usually relies on retrofitting the environment with additional hardware like WiFi or Bluetooth beacons [2, 19, 34, 44, 57]. These solutions have proven highly effective within indoor environments. NavCog3 [2], for example, excels in indoor navigation by employing Bluetooth beacons for precise turn-by-turn guidance. Nakajima and Haruyama [44] exploit the use of visible lights communication technology, utilizing LED lights and a geomagnetic correction method to localize BLV users. However, extending these approaches to support outdoor navigation is not practical. This is particularly evident when considering the substantial initial investment in hardware setup that these systems typically require, making them ill-suited for the larger, unstructured outdoor environment. Furthermore, most of these methods lack the capability to assist with obstacle avoidance and to prevent users from veering off course — both of which are less severe issues indoors compared to outdoors [49]. In contrast, our exploration of using existing street cameras is better suited to address the largely unaddressed challenge of outdoor pedestrian navigation. This approach offers precise localization without requiring supplementary hardware, harnessing street cameras for locating a pedestrian's position. Additionally, it holds the potential to effectively tackle the distinctive challenges posed by the unstructured nature of outdoor environments, including real-time obstacle detection and the interpretation of critical visual cues like street crossing signals.

3 FORMATIVE INTERVIEWS

We conducted semi-structured interviews with six BLV participants to identify BLV pedestrians' challenges in outdoor navigation when using GPS-based systems (RQ1).

3.1 Methods

Participants. We recruited six BLV participants (three males and three females, aged 29–66) by posting on social media platforms and snowball sampling [22]. Table 1 summarises the participants' information. All interviews were conducted over Zoom and lasted about 90 minutes. Participants were compensated $25 for this IRB-approved study.

Procedure. To identify the specific challenges that BLV people face when navigating outdoors, we used a recent critical incident technique (CIT) [17], in which we asked participants to recall and describe a recent time when they navigated outdoor environments using GPS-based assistive technology (AT). For example, we first

Table 1: Self-reported demographics of our participants. Gender information was collected as a free response; our participants identified themselves as female (F) or male (M). Participants rated their assistive technology (AT) familiarity on a scale of 1–5.

PID	Age	Gender	Race	Occupation	Vision ability	Onset	Mobility aid	AT familiarity (1–5)
F1	29	Female	White	Claims expert	Totally blind	At birth	White cane	3: Moderately familiar
F2	61	Female	White	Retired	Light perception only	Age 6	Guide dog	1: Not at all familiar
F3	66	Female	White	Retired	Totally blind	Age 58	Guide dog	2: Slightly familiar
F4	48	Male	Black	Unemployed	Light perception only	Age 32	White cane	3: Moderately familiar
F5	27	Male	Mixed	Unemployed	Totally blind	At birth	White cane	3: Moderately familiar
F6	38	Male	White	AT instructor	Totally blind	At birth	White cane	5: Extremely familiar

asked participants to name the AT they commonly use and then asked them to elaborate on their recent experience of using it: *"So, you mentioned using BlindSquare a lot. When was the last time you used it?"* Then, we initiated a discussion by establishing the scenario for them: *"Now, let's walk through your visit from the office to this restaurant. Suppose, I spotted you at your office. What would I observe? Let's start with you getting out of your office building."* We asked follow-up questions to gain insights into what made the aspects of outdoor navigation challenging and what additional information could help address them.

Interview Analysis. To analyze the interviews, we first transcribed the study sessions in full and then performed thematic analysis [9] involving three members of our research team. Each researcher first independently went through the interview transcripts and used NVivo [47] to create an initial set of codes. Then, all three iterated on the codes together to identify emerging themes.

3.2 Findings: BLV Pedestrians' Challenges in Outdoor Navigation

We found three major themes around challenges that BLV pedestrians face when navigating outdoors using GPS-based systems.

C1: Routing through complex environment layouts. GPS-based systems, such as BlindSquare [41], offer navigation instructions that follow a direct path to the destination from the user's current position, often referred to as "as the crow flies," rather than providing detailed turn-by-turn instructions through a polyline path that guide BLV people through the environment layout. Since *"not everything is organized in the ideal grid-like way"* (F1), participants reported difficulties following the "as the crow flies" instructions, failing to confidently act upon the instructions without any knowledge of the environment layout. This was particularly challenging in complex layouts, as F3 recalled: *"I didn't know if crosswalks were straight or curved or if they were angled. [It was hard] to figure out which way you needed to be to be in the crosswalk."* Many participants cited problems such as making the wrong turns into unexpected "alleyways" (F1, F2, F4) that landed them in dangerous situations with "cars coming through" (F2). Participants cited examples about how these instructions were often inaccurate, causing them to veer off course—a common issue for BLV people in open, outdoor space [49]—and end up in the middle of the streets.

C2: Avoiding unexpected obstacles while using GPS-based systems. BLV people's challenges relating to obstacles during navigation are well researched [50, 52]. However, we found specific nuances in their difficulties, particularly when they rely on their conventional mobility aids in conjunction with GPS-based navigation systems. Participants commonly reported the use of mobility aids like white canes alongside GPS systems for guidance. During this combined navigation process, they encountered difficulties in maintaining their focus on obstacle detection, often resulting in collisions with objects that they would have otherwise detected using their white canes. For instance, F2 shared an incident where they remarked, *"there were traffic cones [and] I tripped over those"* while following directions. Notably, moving obstacles such as pedestrians and cars, as well as temporarily positioned stationary obstacles like triangle sandwich board signs, posed significant challenges for navigation. F4 expressed this sentiment, stating, *"You know how many times I've walked into the sides of cars even though I have the right of way. Drivers have gotten angry, accusing me of scratching their vehicles. It can spoil your day [and make] you feel insecure."*

C3: Crossing street intersections safely. Consistent with prior research [3, 26, 40], our study participants highlighted that crossing streets remained a significant challenge for them. Since GPS-based systems do not help with street-crossing, most participants relied on their auditory senses. They mentioned the practice of listening for vehicular sounds to gauge traffic flow on streets running parallel and perpendicular to their position. This auditory technique helped them assess when it was safe to cross streets. However, participants also reported instances where this method proved inadequate due to external factors: *"yeah, it can be tricky, because [there may be] really loud construction nearby that can definitely throw me off because I'm trying to listen to the traffic"* (F1). Furthermore, their confidence in street-crossing decisions was affected by their inability to ascertain the duration of pedestrian signals and the length of the crosswalk. This uncertainty led to apprehension, as they expressed a fear of becoming stranded mid-crossing, as exemplified by one participant's comment: *"I don't want to be caught in the middle [of the street]"* (F4).

4 THE STREETNAV SYSTEM

StreetNav is a system that explores the concept of repurposing street cameras to support outdoor navigation for BLV pedestrians (RQ2,
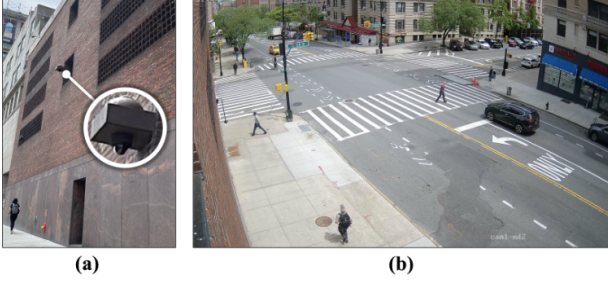
Figure 2: Street camera used for StreetNav's development and evaluation. The COSMOS camera [13, 68] is (a) mounted on the second floor of Columbia University's Mudd building in NYC, and (b) captures the view of a four-way intersection.

RQ3). It provides users precise turn-by-turn navigation instructions to destinations (**C1**), helps prevent veering off track (**C1**), gain awareness of nearby obstacles (**C2**), and assist in crossing streets safely (**C3**). StreetNav enables these navigation affordances through its two main components: (i) *computer vision pipeline*, and (ii) *companion smartphone app*. The computer vision pipeline processes the street camera's video feeds to give BLV pedestrians real-time navigation feedback via the app. Our design and development of StreetNav considers prior work on navigation assistance, functions of traditional mobility aids, and formative interviews with BLV people (Section 3) that identified challenges they face when navigating outdoors using existing GPS-based systems.

The following sections describe StreetNav's technical setup (Section 4.1), the computer vision pipeline (Section 4.2), and the smartphone app's user interface (Section 4.3).

4.1 StreetNav: Technical Setup

Figure 2 shows the street camera we used for developing and evaluating StreetNav. We chose this camera because it faces a four-way street intersection—the most common type of intersection—and is mounted on a building's second floor, offering a typical street-level view of the intersection. The camera is part of the NSF PAWR COSMOS wireless edge-cloud testbed [55, 68]. Anonymized video samples from COSMOS cameras, including the one used in this work, can be found online [13]. StreetNav's computer vision pipeline takes the real-time video feed from the camera as input. For this purpose, we deployed the computer vision pipeline on one of COSMOS' computational servers, which captures the camera's video feed in real time [20, 21]. This server runs Ubuntu 20.04 with an Intel Xeon CPU@2.60GHz and an Nvidia V100 GPU.

StreetNav's two components—the computer vision pipeline and the app—interact with each other via a cloud server, sharing information using the MQTT messaging protocol [43]. Since MQTT is a lightweight messaging protocol, it runs efficiently even in low-bandwidth environments. The computer vision pipeline only sends processed navigation information (e.g., routing instructions, obstacle's category and location) to the app, rather than sending video data. This alleviates any privacy concerns around streaming the video feed to the users and avoids any computational bottlenecks that may happen due to smartphones' limited processing
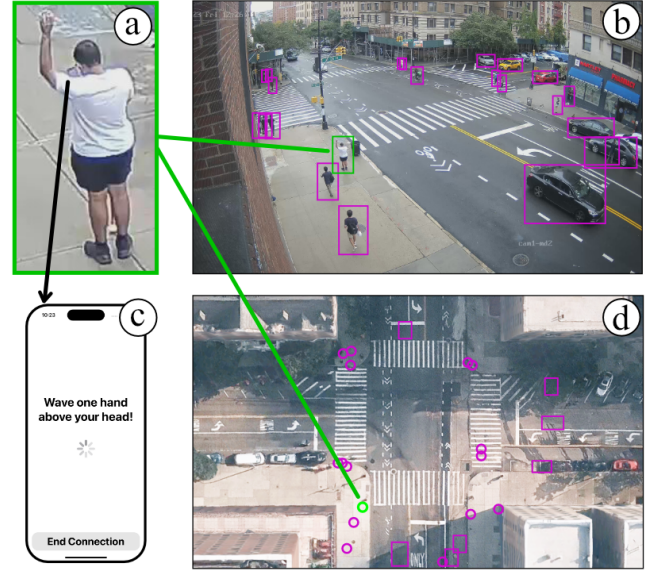

Figure 3: Gesture-based localization for determining a user's position on the map. (a) A study participant (P1) is (c) prompted to wave one hand above their head, enabling the computer vision pipeline to distinguish them from other pedestrians in (b) the camera feed view and (d) the map.

capabilities. The StreetNav app's primary purpose is to act as an interface between the user and the computer vision pipeline. We developed StreetNav's iOS App using Swift [6], enabling us to leverage VoiceOver [7] and other built-in accessibility features.

4.2 StreetNav: Computer Vision Pipeline

StreetNav's computer vision pipeline processes the street camera's video feed in real time to facilitate navigation assistance. It consists of four components: (i) *localizing and tracking the user*: locating user's position on the environment's map; (ii) *planning routes*: generating turn-by-turn navigation instructions from user's current position to destinations; (iii) *identifying obstacles*: predicting potential collisions with other pedestrians, vehicles, and objects (e.g., trash can, pole); and (iv) *recognizing pedestrian signals*: determining when it is safe for pedestrians to cross (walk vs. wait) and calculating the duration of each cycle. Next, we describe the computer vision pipeline's four components in detail.

Localizing and tracking the user. To offer precise navigation assistance, the system must first determine the user's position from the camera view and then project it onto the environment's map. Figure 3d shows the map representation we used, which is a snapshot from Apple Maps' [5] satellite view of the intersection where the camera is deployed.

StreetNav tracks pedestrians from the camera's video feed using Nvidia's DCF-based multi-object tracker [45] and the YOLOv8 object detector [63]. The computer vision pipeline is developed using Nvidia GStreamer plugins [46, 62], enabling hardware-accelerated video processing to achieve real-time tracking. We chose this tracker for its trade-off between real-time performance and robustness to
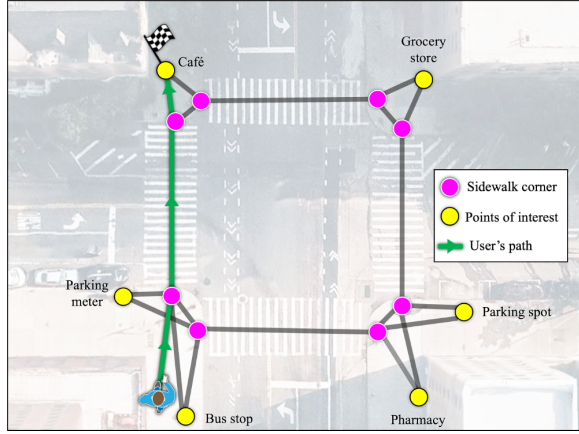
Figure 4: StreetNav's internal graph representation for route planning. The user's current position is added dynamically as a start node to the graph upon choosing a destination. The shortest path, highlighted in green, is then calculated as per this graph representation.

Figure 5: Identifying obstacles in the user's vicinity. (a) A vehicle turning left yields to the BLV pedestrian (detected in purple) crossing the street. (b) StreetNav identifies the obstacles' category and relative location on the map to provide real-time feedback via the app.

occlusions. The tracker detects all pedestrians and assigns them a unique ID. However, the system needs a way to differentiate between the BLV user and other pedestrians.

Figure 3 shows the *gesture-based localization* approach we introduced to address this issue. To connect with the system, BLV pedestrians must wave one hand above their head for 2–3 seconds (Figure 3a), enabling the system to determine the BLV pedestrian's unique tracker ID. We chose this gesture after discussions with several BLV individuals, including our BLV co-author, and most agreed that this single-handed action was both convenient and socially acceptable to them. Moreover, over-the-head gestures such as waving a hand can also be detected when users are not directly facing the street camera.

StreetNav implements the gesture-based localization approach by first creating image crops of all detected pedestrians and then classifying them as 'waving' or 'walking' pedestrians using CLIP [53]. CLIP classifies each pedestrian by computing visual similarity between the pedestrian's image crop and two language prompts: 'person walking' and 'person waving hand.' We experimentally fine-tuned the confidence thresholds and these language prompts. We also tried other action recognition models, such as MMaction2 [12], but found that our CLIP-based approach was much faster and robust to false positives.

Finally, we transformed the user's position on the street camera view (Figure 3b) onto the map (Figure 3d) using a simple feed-forward neural network, trained on data that we manually annotated. The network takes as input the 2D pixel coordinate from the street camera view and outputs the corresponding 2D coordinate on the map. StreetNav continuously tracks the user from the camera feed and transforms its position onto the map.

Planning routes. StreetNav represents routes as a sequence of straight lines on the map connected by waypoints. To plan routes, StreetNav requires that a map of the environment is annotated with waypoints and connections between them. This offline process
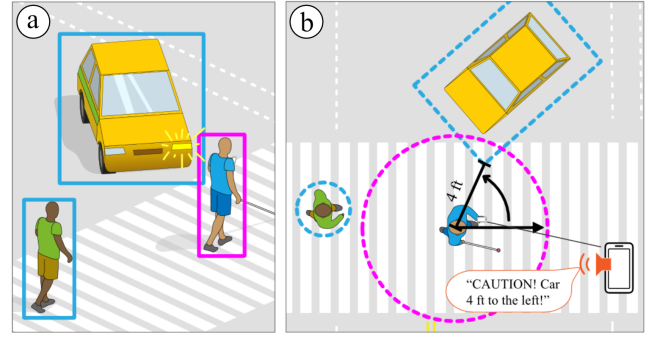
is performed by manually annotating the environment's map, as shown in Figure 4. The administrator marks two types of points on the map: POIs and sidewalk corners. The POIs are potential destinations that users can choose from. The sidewalk corners act as intermediary waypoints en route to the destination. We chose sidewalk corners as waypoints because BLV pedestrians often look for the tactile engravings at sidewalk corners to help orient themselves and transition into crosswalks. Thus, these waypoints blend in well with BLV users' current navigation practices.

Figure 4 shows the internal graph structure that StreetNav uses for planning routes. This graph-based representation of the environment has also been used in prior work on indoor navigation systems [2, 25, 57]. In the graph, nodes correspond to POIs and sidewalk corners, whereas edges correspond to walkable paths. Once the user chooses a destination from the POIs, StreetNav adds the user's current position as a start node to this graph representation and computes the shortest path to the chosen POI using A* algorithm [14]. Figure 4 highlights the shortest path from the user's current position to the chosen destination (café). This route enables StreetNav to guide users to the destination via turn-by-turn instructions.

Identifying obstacles. Prior work on obstacle avoidance developed systems that guide BLV people around obstacles [25, 33]. StreetNav, however, aims to augment BLV pedestrians' awareness of obstacles to help them confidently avoid obstacles using their traditional mobility aids (e.g., white cane) and mobility skills. From our formative interviews, we learned that obstacles that catch BLV users unexpectedly were specifically hard to avoid in outdoor environments (**C2**). Thus, StreetNav provides users with information about the obstacle's category and relative location. This gives BLV users context on the size, shape, and location of an obstacle, enabling them to confidently use their mobility skills around unexpected obstacles.

Figure 5 illustrates how the system identifies obstacles in the user's vicinity. StreetNav's multi-object tracker is used to track other objects and pedestrians. Examples of other objects include cars, bicycles, poles, and trash cans. The computer vision pipeline
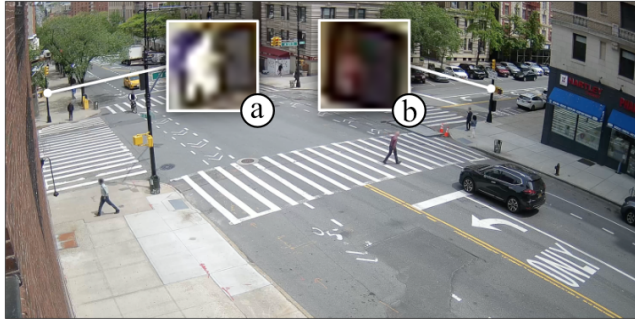
Figure 6: Recognizing pedestrian signal states from the camera's video feed. StreetNav compares the number of white and red pixels in the signal crops to determine its state: (a) *walk* **vs. (b)** *wait.*

then projects the detected objects' positions onto the map. To identify obstacles in the BLV user's vicinity, StreetNav computes the distance and angle between the user and other detected objects with respect to the map (Figure 5b). Any object (or pedestrian) within a fixed radial distance from the BLV user is flagged as an obstacle. Through a series of experiments with our BLV co-author, we found that a 4 foot radius works best for StreetNav to provide users with awareness of obstacles in a timely manner.

Recognizing pedestrian signals*.* To determine the pedestrian signals' state (i.e., *walk* vs. *wait*), we leverage the fact that walk signals are always white, whereas wait signals are always red in color. StreetNav requires the pixel locations of the pedestrian signals in the video feed in order to recognize the signal state. The administrator annotates the video feed image to draw a bounding box around the pedestrian signals' screen. Since the position of pedestrian signals is fixed with respect to the mounted street camera, this process needs to be done only once during setup, along with the map annotation process described earlier.

Figure 6 shows the annotated pedestrian signals in the camera's video feed. StreetNav uses these annotations first to generate image crops of the two signals and then threshold both image crops to filter all red and white pixels. It compares the number of white and red pixels in each crop to identify the signal's state: *walk* (Figure 6a) vs. *wait* (Figure 6b). We experimentally fine-tuned the count thresholds to accurately identify the signal state. Although the two crops are low resolution, this approach still yields accurate results since it distinguishes the state using pixel colors.

Our formative interviews found that BLV pedestrians faced difficulty pacing themselves while crossing streets (**C3**). To address this challenge, StreetNav provides users with information about how much time remains for them to cross. StreetNav's computer vision pipeline computes the time remaining to cross by keeping track of the signal cycles' duration. StreetNav maintains a timer that records the moments when each signal changes its state. After observing a full cycle, StreetNav is able to accurately keep track of both the state and timing of each signal. StreetNav periodically refreshes the timer to adapt to any changes in signal duration that may happen for traffic management reasons.

4.3 StreetNav App: User Interface

The StreetNav iOS app interacts with the computer vision pipeline to allow BLV pedestrians to choose a destination and receive real-time navigation feedback that guides them to it. BLV users first initiate a connection request through the app, which activates the gesture-based localization (Section 4.2) in the computer vision pipeline. The app prompts the user to wave one hand over their head (Figure 3b), enabling the system to begin tracking their precise location on the map (Figure 3d). BLV users can then select a destination from nearby POIs and begin receiving navigation feedback through the app.

Figure 7 shows the StreetNav app's user interface, which uses audiohaptic cues for (i) providing routing instructions, (ii) preventing veering off track, (iii) notifying about nearby obstacles, and (iv) assisting with crossing streets. Upon reaching the destination, the app confirms their arrival. The following sections describe the app's interface in detail.

Providing routing instructions*.* The app conveys routing instructions to the users by first giving an overview of the route and then announcing each instruction, in situ, based on their current location in the environment. Figure 7a shows the app screen with the path overview. Prior work on understanding BLV people's navigation behaviors [1, 24, 30] reveals that BLV people often prepare for their routes before actually walking through them. StreetNav assists them in this preparation by giving an overview of the path before beginning navigation. The path overview consists of several instructions, with each helping them get from one waypoint to the next. BLV users read through the path overview using VoiceOver [7]. Users then tap the 'Start Navigation' button, which announces each instruction when they reach a waypoint. Figure 7b–f shows how the app dynamically updates the next instruction based on the user's location in the environment. Throughout the journey, users can access the path overview and the current navigation instructions on demand via VoiceOver.

Preventing veering off track*.* Figure 8 illustrates the app's feedback for preventing users from veering off track. Given the user's current position, heading, and destination route, StreetNav computes the *direction* and *extent* of veering. To convey *direction* of veering, we used 3D spatialized sound, which plays continuous beeping sounds from the right speaker when users veer to the left (Figure 8a) and from the left speaker when users veer to the right (Figure 8c). Users can follow the direction of the beeping sound to correct for veering. To convey the *extent* of veering. i.e., how severely the user is veering, we render the frequency of beeps to be proportional to the angle between the user's current heading and the route. As users veer away from the correct direction, the frequency of beeps increases; and when they begin to turn towards the correct direction, the frequency of beeps decreases. Users can also leverage the frequency of beeps to determine how to correct for veering, by always moving in the direction where the beeps' frequency reduces. This enables users to correct for veering even without the spatialized sound feedback we used for direction. This eliminates the need to wear headphones to understand spatialized sound.
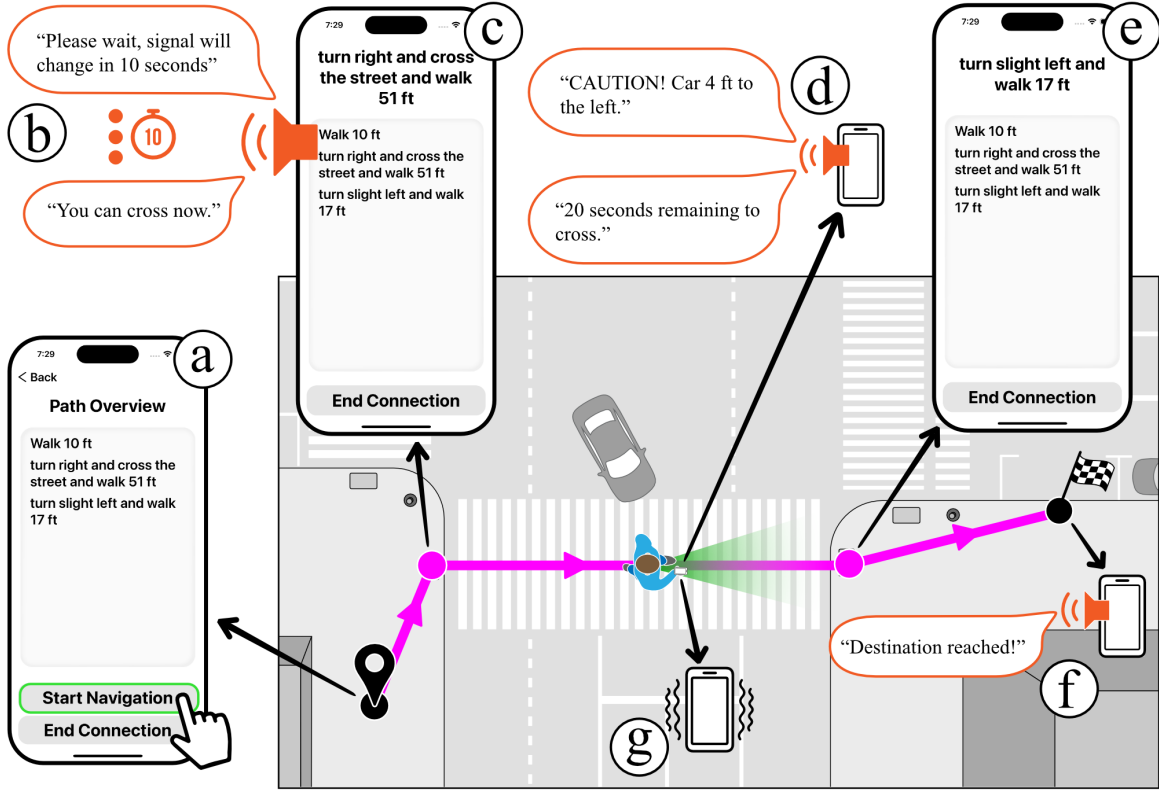
Figure 7: The StreetNav App's user interface. It provides routing instructions to their destination via (a) a path overview and (c, e) real-time feedback that updates their current instruction based on their location. Upon reaching a sidewalk, (b) the app informs the user about when it is safe to cross and (d) how much remains for them to cross over. It also (d) notifies the user of a nearby obstacle's category and relative location to help them avoid it. The app (f) confirms the user's arrival at the destination. Throughout the journey, the app provides (g) continuous audiohaptic feedback to prevent users from veering off track.

We ran pilot experiments to test this feedback mechanism with our BLV co-author. We found that the continuous audio feedback was helpful but also became overwhelming as it forced them to strictly follow StreetNav's route. To address this, we relaxed the veering requirements by introducing a tolerance angle (θ). Figure 8 shows the tolerance angle in green color, which is depicted as a cone centered at the user's current heading. We updated the veering feedback to only play beeping sounds when users veer off in either direction by at least $\theta/2$ degrees. To maintain the continuity of feedback, we chose to render subtle haptic vibrations when users move in the correct direction within the tolerance angle. Within this tolerance angle, the intensity of the haptic vibration increases when users approach the exact correct heading and decreases when they start to veer off. This is similar to how the frequency of beeps increases when users veer away. In this way, the audio feedback acts as negative reinforcement, and the haptic feedback acts as positive reinforcement. Figure 8b illustrates the haptic feedback. We experimentally tuned the tolerance angle, θ, and set its value for our system to $50°$.

To generate the audiohaptic cues, the app receives the user's current position and destination route from the computer vision pipeline. For the user's current heading, we experimented using the user's trajectory to predict their heading using the Kalman filter. This approach, however, yields inaccurate headings due to the noisy tracking data. Thus, we leveraged the smartphone's compass to determine the user's current heading. We offset the compass readings by a fixed value to ensure that its zero coincides with the map's horizontal direction. This enabled us to perform all heading-related computations with respect to the map's frame of reference.

Notifying about nearby obstacles. Figure 7d shows how StreetNav alerts the user of obstacles nearby. The app announces the obstacle's category, distance, and relative location. For example, when a car approaches the user, the app announces: "*Caution! Car, 4 ft. to the left.*" Similar to veering feedback, the relative location is computed using both the computer vision pipeline's outputs and the smartphone's compass reading.

We tried feedback formats with varying granularity to convey the obstacle's relative location. First, we experimented with *clock-faced directions*: "*Car, 4 ft. at 1 o'clock.*" Clock-faced directions are commonly used in many GPS-based systems such as BlindSquare to convey directions. We learned from pilot evaluations with our BLV co-author that this feedback format was too fine-grained, as it took them a few seconds to decode the obstacle's location. This does not
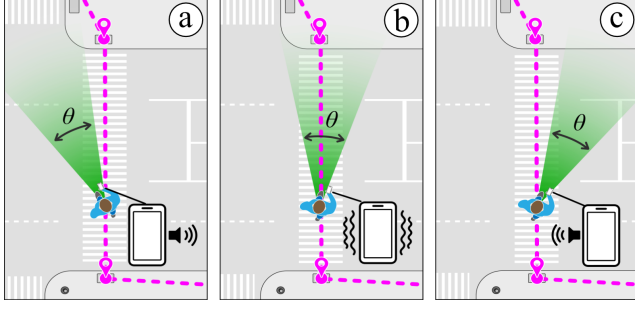
Figure 8: Audiohaptic cues for preventing users from veering off track. Sample user trajectories showing feedback when users (a) veers to the left, (b) do not veer, and (c) veer to the right. When the user's heading coincides with the route to the destination, within a tolerance angle θ (highlighted in green), users receive (b) subtle haptic vibrations to reinforce them. When they veer off the route, outside the tolerance angle θ, they hear spatialized beeping sounds that are rendered from the (a) right speaker when veering left, and from the (c) left speaker when veering right.
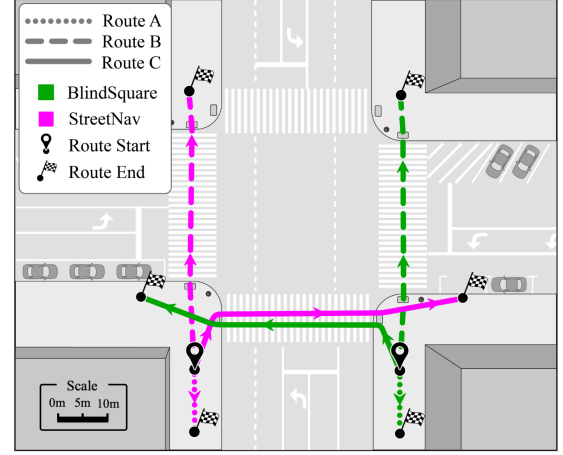

Figure 9: The routes used in the navigation tasks. (A) 12 meters, stationary person to avoid on the sidewalk. (B) 30 meters, cross street, and moving person to avoid on the sidewalk. (C) 38 meters, a $90°$ turn, cross street, and moving person to avoid on the crosswalk. To mitigate learning effects, routes for the two conditions are symmetrically designed, situated on opposite sides of the street.

fare well with moving obstacles, such as pedestrians, that may have already passed the user before they are able to decode the location. Moreover, StreetNav's goal with obstacle awareness is to give users a quick idea that something is nearby them, which they can then use to circumnavigate via their mobility skills. To address this, we tried the more coarse format with just four directions: left, right, front, and back. This was found to give users a quick intimation, compared to the clock-faced directions.

Assisting with crossing streets. The StreetNav app helps users cross streets by informing them *when* to cross and how much time remains before the signal changes.

Figure 7b and Figure 7d illustrate the feedback. Upon reaching a sidewalk corner, the app checks for the signal state recognized by the computer vision pipeline. If the signal is '*wait*' when the user arrives, the app informs the user to wait along with the time remaining before the signal changes. If the signal is '*walk*' when the user arrives, the app informs the user to begin crossing only if the time remaining is sufficient for crossing. For the intersection used in our user studies, this was experimentally found to be 15 seconds. Otherwise, the user is advised to wait for the next cycle. Once the user begins crossing on the '*walk*' signal, the app announces the time remaining for them to cross over. This feedback is repeated at fixed intervals until the user reaches the other sidewalk corner. We experimentally fine-tuned this interval with feedback from our BLV co-author. We tried several intervals, such as 5, 10, and 15 seconds, and found that shorter intervals overwhelmed the users, whereas longer intervals practically would not be repeated enough times to give them meaningful information. We settled on repeating the feedback every 10 seconds for our implementation.

5 USER STUDY

Our user study had three goals, related to RQ2 and RQ3. First, we wanted to evaluate the extent to which StreetNav addressed BLV

pedestrians' challenges in navigating outdoor environments when using existing GPS-based systems. Through our formative interviews (Section 3), we discovered three main challenges: routing through complex environment layouts (**C1**), avoiding unexpected obstacles (**C2**), and crossing street intersections (**C3**). Second, we wanted to analyze BLV pedestrians' experience of navigating outdoors using StreetNav compared to existing GPS-based systems. Third, we wanted to see how participants rank the two navigation systems—StreetNav vs. GPS-based system—in order of their preference for outdoor navigation assistance.

5.1 Study Description

Participants. We recruited eight BLV participants (five males, three females; aged 24–52) by posting to social media platforms and by snowball sampling [22]. Participants identified themselves with a range of racial identities (Asian, Black, White, Latino, and Mixed) and all of them lived in a major city in the US. Participants also had diverse visual abilities, onset of vision impairment, and familiarity with assistive technology (AT) for navigation.

Table 2 summarizes participants' information. All but three participants (P1, P7, and P8) reported themselves as being moderately–extremely experienced with AT for navigation (3+ scores on a 5-point rating scale). Only P3 reported minor hearing loss in both ears and wore hearing aids. All participants except two (P2, P9) used white cane as their primary mobility aid. P2 did not use any mobility aid, while P9 primarily used a guide dog for navigation. The IRB-approved study lasted for about 120 minutes, and participants were compensated $75 for their time.

Experimental Design. In the study, participants completed three navigation tasks at a street intersection in two conditions: (i) Street-Nav and (ii) BlindSquare [41], a popular GPS-based navigation

Table 2: Self-reported demographics of our study participants. Gender information was collected as a free response. Participants rated their familiarity with assistive technology (AT) on a scale of 1–5.

PID	Age	Gender	Occupation	Race	Vision ability	Onset	Mobility aid	AT familiarity (1–5)
P1	24	Male	App developer	Asian	Low vision	Age 19	White cane	2: Slightly familiar
P2	28	Male	Data manager	White	Low vision	At birth	None	3: Moderately familiar
P3	48	Male	Not employed	Black	Totally blind	Age 32	White cane	3: Moderately familiar
P4	46	Female	Social worker	Latino	Totally blind	Age 40	White cane	4: Very familiar
P5	43	Female	Not employed	Asian	Totally blind	At birth	White cane	4: Very familiar
P6	52	Male	Mgmt. analyst	Mixed	Light perception only	Age 9	White cane	5: Extremely familiar
P7	26	Female	Writer	Mixed	Low vision	At birth	White cane	2: Slightly familiar
P8	51	Male	Not employed	Black	Light perception only	Age 26	Guide dog	3: Moderately familiar

app especially designed for BLV people. We evaluated the two systems via their respective iOS apps on an iPhone 14 Pro. Both systems' apps seamlessly integrated with VoiceOver, and all eight participants had a high level of familiarity with using iPhones and VoiceOver, with ratings of 3 or higher on a 5-point scale. During the study, participants continued to use their primary mobility aids, such as white canes and guide dogs, in both conditions. This approach allowed us to make a meaningful comparison between StreetNav and the BLV pedestrians' current methods of outdoor navigation, simulating their usual practice of incorporating GPS-based navigation systems alongside their mobility aids.

Our study followed a within-subjects design, in which participants tested the two navigation systems in a counter-balanced order to minimize potential order-bias and learning effects. In each condition, participants were tasked with completing three distinct navigation challenges, corresponding to three specific routes. Figure 9 illustrates these three navigation routes. We deliberately chose the routes to lie within the street camera's field of view and include a range of difficulty levels for each task: (A) a short route, 12 meters, that involved avoiding a stationary person on the sidewalk, (B) a long route, 30 meters, that involved crossing a street and avoiding a moving person on the sidewalk, and (C) a complex route, 38 meters, that involved making a 90 degree turn, crossing a street, and avoiding a moving person on the crosswalk. For each of these tasks, one of our researchers assumed the role of the obstacle. Notably, none of the participants were familiar with the specific street intersection selected as the study's location.

Given that participants navigated the same intersection in both conditions, the potential for learning effects as a confounding factor was carefully considered. To address this concern, we took deliberate measures by creating distinct routes for each condition. Specifically, we designed the routes in both conditions to be symmetric—rather than being identical—with the starting and ending points of each route strategically positioned on opposite sides of the street intersection, as illustrated in Figure 9. The symmetry of routes ensured that participants encountered the same challenges in both conditions. To ensure participants' safety, the researchers accompanied them at all times during the study, prepared to intervene whenever necessary.

Procedure. We began each study condition by giving a short tutorial of the respective smartphone app for the system. During these

tutorials, participants were taught how to use the app and how to interpret the various audiohaptic cues it offered. To accommodate potential challenges arising from ambient noise at the street intersection, participants were given the option to wear headphones during the study. Only two participants, namely P3 and P5, exercised that option; rest of the participants relied on the smartphone's built-in speaker to hear the audiohaptic cues.

After completing the three navigation tasks for each condition, we administered a questionnaire comprising four distinct parts. These parts were designed to assess participants' experiences around challenges faced by BLV pedestrians in outdoor navigation, specifically addressing the following aspects: routing to destination (**C1**), veering off course (**C1**), avoiding obstacles (**C2**), and crossing streets (**C3**). It included questions about how well each system assisted with the challenges, if at all. Participants rated their experience on a 5-point rating scale, where a rating of "1" indicated "*not at all well,*" and a rating of "5" indicated "*extremely well.*" After each part of the questionnaire, we asked follow-up questions to gain deeper insights into the reasons behind their ratings and their overall experiences.

Following their experience with both navigation systems, participants were asked to complete a post-study questionnaire. This questionnaire required them to rank the two navigation systems in terms of their preference for outdoor navigation. Subsequently, we directed our discussion toward StreetNav, engaging participants in a conversation about potential avenues for improvement. We also inquired about the specific scenarios in which they envision using this system in the future.

In addition to the questionnaires that aimed at capturing participants' subjective experiences, we also gathered system usage logs and video recordings of participants throughout the study. These objective data sources, including usage logs and video recordings, allowed us to perform a comprehensive analysis of participants' actual performance in the navigation tasks. It is worth noting that willingness to be video-recorded was completely voluntary, i.e., did not affect participants' eligibility or compensation. All eight participants still agreed to be video-recorded, providing us with written consent to do so.

Analysis. We report participants' spontaneous comments that best represent their overall opinions, providing further context on the quantitative data we collected during the study. We analyzed the
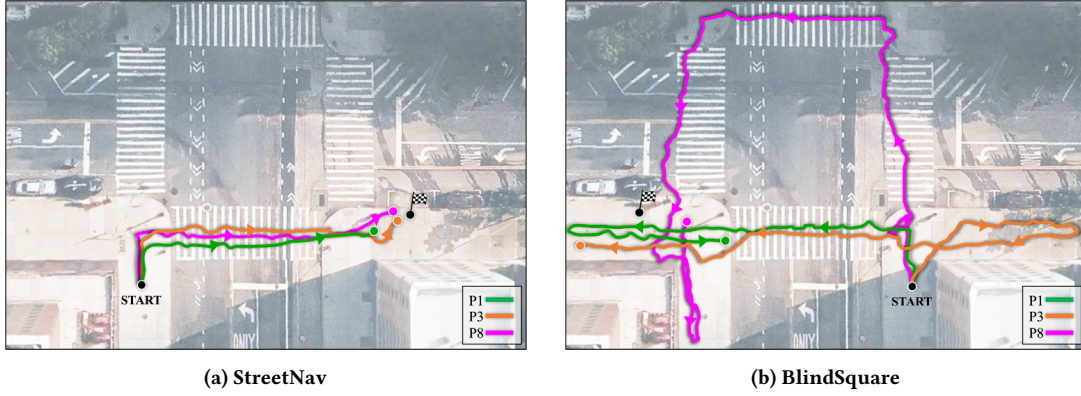
(a) StreetNav

(b) BlindSquare

Figure 10: Comparison of paths traveled by three participants (P1, P3, P8) for route 'C' using (a) StreetNav, and (b) BlindSquare. StreetNav's routing instructions consistently guided participants to the destination via the shortest path. BlindSquare, however, caused participants to take incorrect turns (P1, P3, P8), oscillate back and forth near destinations (P1, P8), and even go around the whole intersection before getting close to the destination (P8).
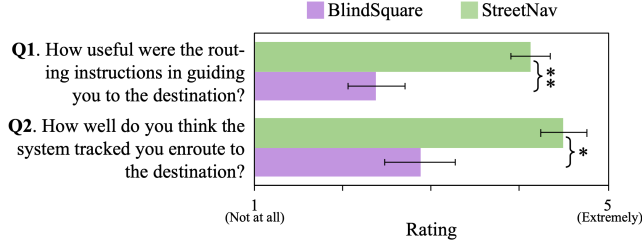

Figure 11: Results for participants' experience with routing to the destination. Participants rated the (1) usefulness of routing instructions, and (2) the system's ability to track them en route to the destination. Participants found StreetNav's turn-by-turn instructions significantly more useful and precise than BlindSquare's "as the crow flies"-style routing instructions. Pairwise significance is depicted for $p < 0.01$ ($*$) and $p < 0.05$ ($**$). The error bars indicate standard error.

transcripts for participants' quotes and grouped them according to the (i) questionnaire's four parts: routing to destination, veering off course, avoiding obstacles, and crossing streets; (ii) overall satisfaction and ranking preferences, and (iii) how users' individual experiences influenced their preferences.

5.2 Results

Our results reveal that StreetNav helped participants reach their destinations with more precision, gain awareness of obstacles, reduce veering off course, and confidently cross streets. For the statistic analysis of each measure, we first conducted a Kolmogorov-Smirnov test to determine if the data was parametric or non-parametric. Then, when comparing between the two conditions, we used a paired t-test when the data was parametric. In addition to quantitative measures, we conducted a detailed analysis of video recordings, manually annotating the routes participants took during the study. We provide these metrics to offer additional insights into participants' performance across both experimental conditions.

Routing to Destination. Figure 11 shows participants' average rating for their experience following routes to the destination in each condition. The mean (± std. dev.) rating for participants' perceived usefulness of the routing instructions in guiding them to the destination was 4.13 (±0.64) for StreetNav and 2.38 (±0.91) for BlindSquare. The condition had a significant main effect ($p = 0.014$) on participants' experience reaching destinations with the routing instructions. The mean (± std. dev.) rating for participants' experience with the system's ability to track them was 4.50 (±0.76) for StreetNav and 2.88 (±1.13) for BlindSquare. The condition had a significant main effect ($p = 0.001$) on participants' perception of how well the system tracked them en route to the destination. This indicates that participants found StreetNav more useful than BlindSquare for guiding them to the destination.

Figure 10 illustrates our analysis of the video recordings, plotting the typical paths taken by participants in the third route across both conditions. We computed various metrics from their paths, that provide insights into participants' self-reported ratings.

We found that when using BlindSquare, participants covered greater distances to reach the same destinations compared to when using StreetNav. On average, participants traveled a distance approximately 2.1 times longer than the shortest route when relying on BlindSquare. In contrast, when using StreetNav, they covered a distance of only about 1.1 times the shortest route to their destination. This represents a 51% reduction in the unnecessary distance traveled with StreetNav in comparison to BlindSquare. Figure 10b shows how participants using BlindSquare often exhibited an oscillatory pattern near their destinations (P1, P8) before eventually reaching close to them.

Additionally, StreetNav's routing instructions displayed a notably higher level of precision, guiding participants to their destinations with 2.9 times greater accuracy than BlindSquare. Figure 10 clearly shows this trend for the third route. On average, across the three study routes, participants using StreetNav concluded their journeys within a tighter radius of 12.53 feet from their intended destination. In contrast, participants relying on BlindSquare concluded their journeys within a radius of 35.94 feet from their
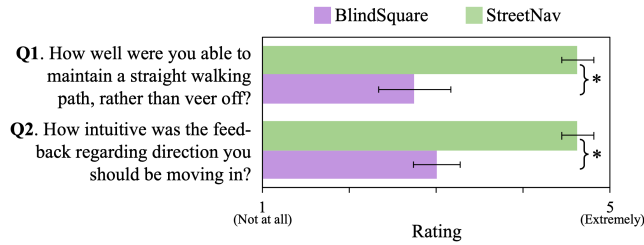
Figure 12: Results for participants' perceived ability to prevent veering off path. Participants rated their ability to (1) maintain a straight walking path, and (2) intuitiveness of the feedback regarding direction they should be moving in; on a scale of 1–5. StreetNav'saudiohaptic feedback was significantly more intuitive than BlindSquare's in preventing participants from veer off path. Pairwise significance is depicted for $p < 0.01$ (∗). The error bars indicate standard error.

intended destination. Two study participants, P4 and P5, even refused to navigate to the destination in two of the three tasks with BlindSquare. This was primarily attributed to BlindSquare's low precision in tracking the participants and often guiding them to take incorrect turns. Figure 10b highlights how BlindSquare caused P8 to go around the intersection before finally getting close the destination.

Participants preferred StreetNav over BlindSquare for its audiohaptic cues for turn-by-turn navigation instructions, which they found to be more useful and precise than BlindSquare's "as the crow flies"-style clock face and distance-based instructions. P3's comment encapsulates this sentiment:

> "When it's time for me to turn right and walk a certain distance, [StreetNav] is very, very, very precise." –**P3**

Although all participants preferred StreetNav's routing feedback over BlindSquare's, distinct patterns emerged in their preference and utilization of these cues. StreetNav delivers a combination of audiohaptic and speech feedback for routing, and participants adopted varying strategies for utilizing this feedback. Some individuals placed greater reliance on the veering haptic feedback as their primary directional guide, while reserving speech feedback as a fallback option. Conversely, some participants prioritized the speech feedback, assigning it a higher level of importance in their navigation process compared to audio-haptic cues.

Veering Prevention. Figure 12 shows participants' average rating for their perceived ability to (1) maintain a straight walking path, i.e., prevent veering off course, and (2) intuitiveness of the feedback they received regarding direction to move in. The mean (± std. dev.) rating of participants' perceived ability to maintain a straight walking path with StreetNav was 4.63 (±0.52) and with BlindSquare was 2.75 (±1.17). The condition had a significant main effect ($p = 0.001$) on participants' perceived ability to prevent veering off course. The mean (± std. dev.) rating for intuitiveness of the feedback that helped them know which direction to move in was 4.63 (±0.52) for StreetNav and 3.00 (±0.76) for BlindSquare. The condition had a significant main effect ($p = 0.006$) on intuitiveness of feedback that helped participants prevent veering off path.

Our examination of the video recordings aligns closely with participants' ratings. It reveals that StreetNav minimized participants' deviations from the shortest path to the destinations in comparison to BlindSquare. Over the course of the three routes, participants displayed an average deviation from shortest path, that was reduced by 53% when using StreetNav as opposed to BlindSquare.

With BlindSquare, many participants reported difficulty maintaining awareness of their surroundings, including both obstacles and navigation direction, which frequently led to deviations from their intended paths. For instance, P2 reported challenges in maintaining their orientation with the need to avoid obstacles:

> "[BlindSquare] basically demanded me to keep track of my orientation as I was moving, which is pretty difficult to do when you're also trying to keep other things in mind, like not bumping into things." –**P6**

In contrast, StreetNav effectively addressed this challenge by providing continuous audiohaptic feedback for maintaining a straight walking path, instilling a sense of confidence in participants. P3, who tested StreetNav before BlindSquare, reflected on their desire for a similar continuous feedback mechanism within BlindSquare, akin to the experience they had with StreetNav:

> "[with BlindSquare] even though I couldn't see the phone screen, my eyes actually went towards where I'm holding the screen. It is almost as if on a subconscious level, I was trying to get more feedback. With [StreetNav] I had enough feedback." –**P3**

Many participants appreciated StreetNav's choice of haptic feedback for veering. Some participants envisioned the haptic feedback to be especially useful in environments with complex layouts:

> "In the [areas] where the streets are very slanted and confusing. I think haptic feedback will be especially helpful." –**P5**

Other participants highlighted the advantage of haptic feedback in noisy environments where audio and speech feedback might be less effective.

However, both P4 and P6 exclaimed that StreetNav's haptic feedback would only work well when holding the phone in their hands. This meant that hands-free operation of the app may not be possible, which is important for BLV people since one of their hands is always occupied by the white cane. P4 proposed integrating the app with their smartwatch for rendering the haptic feedback to enable hands-free operation.

Obstacle Awareness. Figure 13 shows participants' average rating for their perceived awareness of obstacles across the two conditions. Specifically, participants rated their ability to (1) avoid obstacles, (2) identify its category (e.g., person, bicycle, trash can), and (3) determine its relative location. The mean (± std. dev.) rating for participants' perceived ability to avoid obstacles was 4.38 (±0.74) for StreetNav and 2.88 (±0.99) for BlindSquare, to identify its category was 4.50 (±0.76) for StreetNav and 3.13 (±1.46) for BlindSquare, and to determine obstacle's relative location was 4.13 (±0.64) for StreetNav and 2.88 (±1.25) for BlindSquare. A paired t-test revealed that the condition had a significant main effect on participants' perceived ability to avoid obstacles ($p = 0.030$), identify its category ($p = 0.037$), and relative location ($p = 0.004$). This suggests that
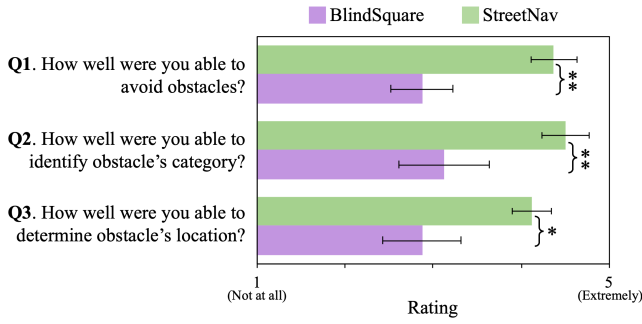
Figure 13: Results for participants' perceived obstacle awareness. Participants rated their ability to (1) avoid obstacles, (2) identify its category (e.g., person, bicycle), and (3) determine its relative location; on a scale of 1–5. StreetNav significantly improved participants' awareness of nearby obstacles during navigation. Pairwise significance is depicted for $p < 0.01$ (∗) and $p < 0.05$ (∗∗). The error bars indicate standard error.

StreetNav offered users a heightened awareness of nearby obstacles compared to the baseline condition of BlindSquare.

With StreetNav, participants had the option to use obstacle avoidance audio feedback in conjunction with their conventional mobility aids. However, in the case of BlindSquare, the system itself did not offer any obstacle-related information. Consequently, participants primarily relied on their traditional mobility aids in this condition, as is typical when using GPS-based systems. Our analysis of the video recordings found that in both experimental conditions, participants encountered no instances of being severely hindered by obstacles. Instead, they adeptly navigated around obstacles with the assistance of their white canes or guide dogs.

Although participants generally had a positive perception of obstacle avoidance when using StreetNav, their opinions on the utility of obstacle awareness information varied. Some participants found this information beneficial, emphasizing its role in preventing "*awkward bumping into people*" (**P2**) and boosting their confidence, resulting in greater "*speed in terms of walking*" (**P3**). Conversely, participants who felt confident avoiding obstacles with their mobility aids regarded StreetNav's obstacle information to be extraneous. P8 also expressed concerns about the potential information overload it could cause in dense urban areas:

> "*To know where people are, is a bit of overkill. If you turn this thing on in Times Square, it would have your head go upside down.*" –**P8**

Many participants proposed an alternative use case for StreetNav's obstacle awareness information, highlighting its potential for providing insights into their surroundings. They suggested that this information could unlock environmental affordances, including the identification of accessible light signals and available benches for resting: "*knowing there was a bench was top-notch for me*" (**P8**). Therefore, StreetNav's obstacle awareness information served a dual purpose, aiding in both obstacle avoidance and environmental awareness, allowing users to "*know what's around*"(**P8**) them.

Crossing Streets. Figure 14 shows participants' average rating for their perceived comfort in crossing streets. The mean (± std. dev.)
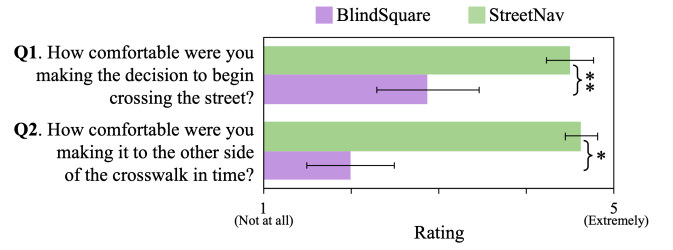

Figure 14: Results for participants' perceived comfort in crossing streets. Participants rated their perceived comfort in (1) making the decision on when to begin crossing the street, and in (2) pacing themselves when crossing. Participants were significantly more comfortable crossing streets with StreetNav in comparison to BlindSquare. Pairwise significance is depicted for $p < 0.01$ (∗) and $p < 0.05$ (∗∗). The error bars indicate standard error.

rating of participants' perceived comfort in making the decision on when to begin crossing the street was 4.50 (±0.76) for StreetNav and 2.88 (±1.64) for BlindSquare. The mean (± std. dev.) rating of participants' perceived comfort in safely making it through the crosswalk and reach the other end was 4.63 (±0.52) for StreetNav and 2.00 (±1.41) for BlindSquare. A paired t-test showed that the condition had a significant main effect on participants' comfort in beginning to cross streets ($p = 0.029$) and in safely making it to the other side ($p = 0.001$).

As BlindSquare does not provide feedback specifically for crossing streets, participants reported relying on their auditory senses, listening for the surge of parallel traffic. However, during the semi-structured interviews, some participants highlighted challenging scenarios that can make this strategy less reliable. P4, for instance, pointed out that ironically, less traffic can complicate street crossings:

> "*I don't always know when to cross because it's so quiet. And sometimes two, three light cycles go by, and I'm just standing there.*" –**P4**

This issue has been exacerbated by the presence of electric cars, which are difficult to hear due to their quiet motors. For P3, their hearing impairments made it challenging to listen for traffic. Thus, most participants appreciated StreetNav's ability to assist with crossing streets:

> "*When it's quiet, I would cross. But now with hybrid cars, it's not safe to do that. [StreetNav] app telling you which street light is coming on is really helpful.*" –**P7**

Participants made decisions to cross the streets by combining StreetNav's feedback with their auditory senses. Many participants emphasized that having information about the time remaining to cross significantly boosted their confidence, especially when this information aligned with the sounds of traffic: "*I thought it was great because I could tell that it matched up*" (**P8**). This alignment between the provided information and their sensory perception inspired confidence in participants:

> "*Relying on my senses alone feels like a gamble about 90 percent of the time, so a system like [StreetNav] that*

accurately displays the amount of time I have to cross the street is great." —**P2**

5.3 Forced Ranking Results

All eight participants unanimously chose StreetNav over Blind-Square as their preferred navigation assistance system. We asked participants to also rank their preferred type of routing instructions. All eight participants strongly preferred StreetNav's turn-by-turn routing instructions compared to BlindSquare's "as the crow flies," direction and distance-style routing instructions.

In the semi-structured interview, participants were asked to elaborate on their rankings. Participants pointed out multiple navigation gaps in BlindSquare, with P2 summarizing participants' sentiment:

> *"If you're only getting somebody 90 percent of the way there, you're not really achieving what I would consider to be the prime functionality of the system."* —**P2**

In contrast, participants praised StreetNav for its precision and real-time feedback, emphasizing the importance of granular and holistic information to support all facets of navigation. However, participants did acknowledge occasional "glitchiness" (**P7**) with StreetNav, which occurred when they moved out of the camera's field of view or were occluded by other pedestrians or vehicles, resulting in lost tracking. Nevertheless, participants still regarded StreetNav as a significant enhancement to their typical navigation experiences, expressing increased confidence in exploring unfamiliar outdoor environments in the future.

> *"It would encourage me to do things that I would not usually... It would make me more confident about going out by myself."* —**P4**

Participants also appreciated StreetNav's ability to identify them in near real-time:

> *"What I found very interesting about the connection part is how quickly it identifies where I am, as soon as I waved my hand, it senses me."* —**P3**

Participants also provided suggestions for improving StreetNav. Some participants wanted a hands-free version that would allow them to hold a white cane in one hand while keeping the other free. Additionally, while they found the gesture of waving hands for connecting with the system socially acceptable, they acknowledged that it might be perceived as somewhat awkward by others in the street.

> *"[Waving a hand] may seem kind of weird to people who don't understand what is going on. But for me personally, I have no issue."* —**P3**

While the gesture-based localization was generally accurate, there were instances where other pedestrians were incorrectly detected as the study participant. On average, the gesture-based localization worked accurately over 90% of the time.

5.4 How Individual Experiences Influenced Participants' Preferences

Throughout the study, participants offered feedback based on their unique backgrounds. We observed distinct patterns in their preferences, affected by their (i) onset of vision impairment, (ii) level of vision impairment, and (iii) familiarity with assistive technology.

Onset of vision impairment. Participants with early onset blindness preferred nuanced, concise feedback with an emphasis on environmental awareness. They used the system as an additional data point without complete reliance. In contrast, participants with late onset blindness trusted the system more and relied heavily on its feedback.

Level of vision impairment. Totally blind participants appreciated the veering feedback, while low-vision users, who had more visual information, relied on their senses and did not need as much assistance with veering. Low-vision participants appreciated the street crossing feedback rather than trying to glean information from pedestrian signals across the street. Totally blind participants relied more on listening for parallel traffic—their usual mode of operation—and used StreetNav's street-crossing feedback as a confirmation.

Familiarity with assistive technology (AT). We noticed that participants who commonly use AT for navigation quickly adapted to StreetNav, while those with less experience hesitated in trusting StreetNav's feedback and had a slightly steeper learning curve. Still, all participants mentioned feeling more comfortable with Street-Nav as the study progressed. Both groups also expressed increased confidence in exploring new areas with StreetNav.

6 DISCUSSION

Our goal with StreetNav was to explore the idea of repurposing street cameras to support precise outdoor navigation for BLV pedestrians. We reflect upon our findings to discuss how street camera-based systems might be deployed at scale, implications of a street camera-based navigation approach for existing GPS-based navigation systems, and the affordances enabled by precise, real-time outdoor navigation assistance.

Deploying street camera-based navigation systems at scale. StreetNav demonstrates that street cameras have the potential to be repurposed for supporting precise outdoor navigation for BLV pedestrians. Our study results show that street camera-based navigation systems can guide users to their destination more precisely and prevent them from veering off course (Figure 10). Our results also show that street camera-based systems can support real-time, scene-aware assistance by notifying users of nearby obstacles (Figure 13) and giving information about when to cross streets (Figure 14). These benefits of a street camera-based approach, over existing GPS-based systems, underscores the need for deploying such systems at scale. Although our system, StreetNav, was deployed at a single intersection, we learned insights on potential challenges and considerations that must be addressed to deploy street camera-based systems at scale.

Several internal and external factors need to be considered before street cameras can be effectively leveraged to support blind

navigation at scale. External factors, including lighting conditions and occlusions on the street, may affect system performance. For instance, we noticed that StreetNav's ability to track pedestrians was affected severely in low-light conditions (e.g., at night) and by occlusions due to the presence of large vehicles (e.g., trucks, buses) and the installation of scaffoldings for construction. Such challenges affect the reliability of street camera-based systems and may limit its operational hours. Internal factors, including the positioning of cameras, their field of view, and variability in resolution, may affect the extent to which such systems can promise precise navigation assistance. For instance, the visibility of the pedestrian signals from the camera feed could affect how much such systems can assist users with crossing streets. With StreetNav, we observed a drop in tracking accuracy as individuals and objects moved further away from the camera.

Therefore, deploying street camera-based systems at scale would require future work to investigate the extent to which both external factors (e.g., lighting, occlusions) and internal factors (e.g., camera resolution) affect system performance and reliability. To address some of the technical limitations around tracking performance and field of view limitations, future research could explore integrating multiple cameras at various elevations and viewing angles. Prior work on robot navigation has explored the fusion of multiple cameras to improve tracking performance [10, 48, 51]. Future work could also explore an ecosystem of accessible street cameras that can share information to automatically manage hand-offs across street intersections, providing users with a seamless experience beyond a single street intersection. Such ecosystems, which span beyond one intersection to a whole district or city, could enable new affordances, such as automatically sensing pedestrian traffic to inform traffic signals and vice versa [35].

Implications for GPS-based navigation systems. When cameras are available, and conditions align favorably, street camera-based systems offer BLV individuals a valuable source of fine-grained, high-precision information, significantly enhancing their navigational experience and environmental awareness. These capabilities are currently beyond the reach of conventional GPS-based systems. All eight study participants unanimously chose StreetNav over BlindSquare as their preferred navigation system due to its precise, scene-aware navigation assistance (Section 5.3). However, it's important to acknowledge that street camera-based systems have their own set of limitations. The widespread availability of street cameras is not yet a reality, and ideal conditions may not always be met for their effective use. In contrast, GPS-based systems, while lacking in precision and environmental awareness, are universally accessible and resilient in varying conditions, including low light. A harmonious integration of these two approaches is a promising solution. Users can tap into street-camera information when conditions permit, seamlessly transitioning to GPS data when necessary. This can be facilitated through sensor fusion or information hand-offs, creating a synergy that ensures a smooth and reliable navigational experience. Future approaches could explore how these two systems can effectively complement each other, addressing their respective limitations and enhancing overall performance.

Affordances of precise outdoor navigation assistance for BLV people. Previous research in indoor navigation has demonstrated the advantages of accurately pinpointing users' locations [2, 34, 57] and providing scene-aware navigational information [25, 33]. However, achieving such precision has remained a challenge in outdoor environments, primarily due to the limited accuracy of GPS technology [23]. StreetNav's approach of leveraging existing street cameras demonstrates that precise outdoor navigation support for BLV pedestrians is possible. Our study reveals the advantages of precise, fine-grained navigation for BLV individuals. These benefits include a substantial reduction in instances of veering and routing errors, such as deviation from the shortest path or missing intended destinations, as well as augmented environmental awareness.

StreetNav offered our participants a glimpse into the potential of precise outdoor navigation. Several participants desired even greater precision, including the ability to discern the exact number of steps remaining before reaching a crosswalk's curb. Future research could delve into exploring how to best deliver such granular feedback to BLV users, alongside the necessary technological advancements needed to achieve this level of precision. These advantages, as our findings suggest, extend beyond merely improving navigation performance. Participants shared insights into how precise navigation could enhance their independence when navigating outdoors. It could empower BLV people to venture outdoors more frequently, unlocking new travel opportunities, as exemplified by P3's newfound confidence in using public transportation with StreetNav-like systems:

> "I don't really use the city buses, except if I'm with somebody, but [StreetNav] would make me want to get up, go outside, and walk to the bus stop." –**P3**

This newfound confidence is particularly noteworthy, considering the unpredictable nature of outdoor environments. Future research could explore new affordances that street camera-based systems can enable for people, in general.

7 LIMITATIONS

Our work revealed valuable insights into the benefits and effectiveness of a new approach that uses existing street cameras for outdoor navigation assistance. At the same time, we acknowledge that our work has several limitations.

StreetNav was developed using a single street camera and tested at a single street intersection. This approach means that there might be other technical hurdles and design considerations we didn't encounter due to the constraints of this setup. Future research could expand upon our design and investigate how street camera-based systems can adapt to different environments and challenges. Furthermore, to ensure the safety of participants and to fit the user study within a 120-minute timeframe, we designed the study routes to be less complex and dangerous. Real-world outdoor environments can vary significantly from one part of a city, state, or country to another. Our study location may not fully capture the diversity of scenarios BLV individuals encounter when navigating outdoors. Lastly, it's important to note that our study sample consisted of only eight BLV individuals. While their insights are valuable, their preferences for outdoor navigation may not represent the broader BLV community's perspectives. StreetNav was developed in response

to the challenges identified in our formative study, but there could be additional challenges and design possibilities that we haven't explored. Future research should consider a more extensive and diverse participant pool to gain a more comprehensive understanding of the needs and preferences within the BLV community.

8 CONCLUSION

We explored the idea of leveraging existing street cameras to support precise outdoor navigation for BLV pedestrians. Our resulting system, StreetNav, addresses BLV people's challenges in outdoor navigation when using GPS-based systems. Our user evaluation revealed StreetNav's potential to guide users to their destination more precisely than existing GPS-based systems. It also demonstrated its ability to offer real-time, context-aware navigation assistance, aiding in obstacle avoidance and safe street crossings. However, we also uncovered various considerations for street camera-based navigation systems, including challenges and opportunities in deploying such systems at scale. These challenges pave the way for future research to enhance the robustness and reliability of street camera-based navigation solutions. Our work highlights the untapped potential of embedding accessibility directly into urban infrastructure by leveraging existing resources, such as street cameras. We envision a future where such systems seamlessly integrate into urban environments, providing BLV individuals with safe and precise navigation capabilities, ultimately empowering them to confidently navigate their surroundings.

REFERENCES

[1] Nazatul Naquiah Abd Hamid and Alistair D.N. Edwards. 2013. Facilitating route learning using interactive audio-tactile maps for blind and visually impaired people. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*. ACM Press, Paris, France, 37. https://doi.org/10.1145/2468356.2468364

[2] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. 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)*. Association for Computing Machinery, New York, NY, USA, 90–99. https://doi.org/10.1145/2935334.2935361

[3] Dragan Ahmetovic, Roberto Manduchi, James M. Coughlan, and Sergio Mascetti. 2017. Mind Your Crossings: Mining GIS Imagery for Crosswalk Localization. *ACM Transactions on Accessible Computing* 9, 4 (Dec. 2017), 1–25. https://doi.org/10.1145/3046790

[4] Amnesty International. 2021. Surveillance City: NYPD can use more than 15,000 cameras to track people using facial recognition in Manhattan, Bronx and Brooklyn. https://www.amnesty.org/en/latest/news/2021/06/scale-new-york-police-facial-recognition-revealed/

[5] Apple Inc. 2023. Apple Maps. https://www.apple.com/maps/

[6] Apple Inc. 2023. Swift. https://developer.apple.com/swift/

[7] Apple Inc. 2023. VoiceOver. https://www.apple.com/accessibility/vision/

[8] Mauro Avila and Limin Zeng. 2017. A Survey of Outdoor Travel for Visually Impaired People Who Live in Latin-American Region. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '17)*. Association for Computing Machinery, New York, NY, USA, 9–12. https://doi.org/10.1145/3056540.3064953

[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[10] Wen-Chung Chang, Chia-Hung Wu, Wen-Ting Luo, and Huan-Chen Ling. 2013. Mobile robot navigation and control with monocular surveillance cameras. In *2013 CACS International Automatic Control Conference (CACS)*. 192–197. https://doi.org/10.1109/CACS.2013.6734131

[11] Coco Feng. 2019. China the most surveilled nation? The US has the largest number of CCTV cameras per capita. https://www.scmp.com/tech/gear/article/3040974/china-most-surveilled-nation-us-has-largest-number-cctv-cameras-capita

[12] MMAction2 Contributors. 2020. OpenMMLab's Next Generation Video Understanding Toolbox and Benchmark. https://github.com/open-mmlab/mmaction2

[13] COSMOS Project. 2023. Hardware: Cameras. https://wiki.cosmos-lab.org/wiki/Hardware/Cameras

[14] František Duchoň, Andrej Babinec, Martin Kajan, Peter Beňo, Martin Florek, Tomáš Fico, and Ladislav Jurišica. 2014. Path Planning with Modified a Star Algorithm for a Mobile Robot. *Procedia Engineering* 96 (2014), 59–69. https://doi.org/10.1016/j.proeng.2014.12.098

[15] Ping-Jung Duh, Yu-Cheng Sung, Liang-Yu Fan Chiang, Yung-Ju Chang, and Kuan-Wen Chen. 2020. V-eye: A vision-based navigation system for the visually impaired. *IEEE Transactions on Multimedia* 23 (2020), 1567–1580.

[16] 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 (ASSETS '14)*. Association for Computing Machinery, New York, NY, USA, 323–324. https://doi.org/10.1145/2661334.2661344

[17] John C. Flanagan. 1954. The critical incident technique. *Psychological Bulletin* 51, 4 (1954), 327–358.

[18] Frank Hersey. 2017. China to have 626 million surveillance cameras within 3 years. https://technode.com/2017/11/22/china-to-have-626-million-surveillance-cameras-within-3-years/

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

[20] Mahshid Ghasemi, Sofia Kleisarchaki, Thomas Calmant, Jiawei Lu, Shivam Ojha, Zoran Kostic, Levent Gürgen, Gil Zussman, and Javad Ghaderi. 2023. Real-time multi-camera analytics for traffic information extraction and visualization. In *IEEE PerCom Demos*.

[21] Mahshid Ghasemi, Zhengye Yang, Mingfei Sun, Hongzhe Ye, Zihao Xiong, Javad Ghaderi, Zoran Kostic, and Gil Zussman. 2023. Video-based social distancing: evaluation in the COSMOS testbed. *IEEE IoT J.* (2023).

[22] Leo A. Goodman. 1961. Snowball Sampling. *The Annals of Mathematical Statistics* 32, 1 (1961), 148–170. https://www.jstor.org/stable/2237615 Publisher: Institute of Mathematical Statistics.

[23] GPS.gov. 2022. GPS Accuracy. https://www.gps.gov/systems/gps/performance/accuracy/

[24] João Guerreiro, Daisuke Sato, Dragan Ahmetovic, Eshed Ohn-Bar, Kris M. Kitani, and Chieko Asakawa. 2020. Virtual navigation for blind people: Transferring route knowledge to the real-World. *International Journal of Human-Computer Studies* 135 (March 2020), 102369. https://doi.org/10.1016/j.ijhcs.2019.102369

[25] João Guerreiro, Daisuke Sato, Saki Asakawa, Huixu Dong, Kris M. Kitani, and Chieko Asakawa. 2019. CaBot: Designing and Evaluating an Autonomous Navigation Robot for Blind People. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 68–82. https://doi.org/10.1145/3308561.3353771

[26] Richard Guy and Khai Truong. 2012. CrossingGuard: exploring information content in navigation aids for visually impaired pedestrians. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 405–414. https://doi.org/10.1145/2207676.2207733

[27] David L Harkey, Daniel L Carter, Janet M Barlow, and Billie Louise Bentzen. 2007. Accessible pedestrian signals: A guide to best practices. *National Cooperative Highway Research Program, Contractor's Guide for NCHRP Project* (2007).

[28] Microsoft Inc. 2018. Microsoft Soundscape - Microsoft Research. https://www.microsoft.com/en-us/research/product/soundscape/. (2018).

[29] Gaurav Jain, Basel Hindi, Mingyu Xie, Zihao Zhang, Koushik Srinivasula, Mahshid Ghasemi, Daniel Weiner, Xin Yi Therese Xu, Sophie Ana Paris, Chloe Tedjo, Josh Bassin, Michael Malcolm, Mehmet Turkcan, Javad Ghaderi, Zoran Kostic, Gil Zussman, and Brian A. Smith. 2023. Towards Street Camera-based Outdoor Navigation for Blind Pedestrians. In *Proceedings of the 25th International ACM*

SIGACCESS Conference on Computers and Accessibility (ASSETS '23). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3517428.3550394

[30] Gaurav Jain, Yuanyang Teng, Dong Heon Cho, Yunhao Xing, Maryam Aziz, and Brian A. Smith. 2023. "I Want to Figure Things Out": Supporting Exploration in Navigation for People with Visual Impairments. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (April 2023), 63:1–63:28. https://doi.org/10.1145/3579496

[31] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Dragan Ahmetovic, Valeria Alampi, Hironobu Takagi, and Chieko Asakawa. 2018. Insights on Assistive Orientation and Mobility of People with Visual Impairment Based on Large-Scale Longitudinal Data. *ACM Transactions on Accessible Computing* 11, 1 (April 2018), 1–28. https://doi.org/10.1145/3178853

[32] Robert K Katzschmann, Brandon Araki, and Daniela Rus. 2018. Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 3 (2018).

[33] Seita Kayukawa, Keita Higuchi, João Guerreiro, Shigeo Morishima, Yoichi Sato, Kris Kitani, and Chieko Asakawa. 2019. BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19.* ACM Press, Glasgow, Scotland Uk, 1–12. https://doi.org/10.1145/3290605.3300282

[34] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. 2016. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing (SAC '16).* Association for Computing Machinery, New York, NY, USA, 604–611. https://doi.org/10.1145/2851613.2851716

[35] Zoran Kostic, Alex Angus, Zhengye Yang, Zhuoxu Duan, Ivan Seskar, Gil Zussman, and Dipankar Raychaudhuri. 2022. Smart City Intersections: Intelligence Nodes for Future Metropolises. *Computer* 55, 12 (Dec. 2022), 74–85. https://doi.org/10.1109/MC.2022.3206273

[36] Laura Griffin. 2020. Surveillance Cameras Are Everywhere. And They're Only Going To Get More Ubiquitous. https://crimereads.com/surveillance-cameras-are-everywhere-and-theyre-only-going-to-get-more-ubiquitous/

[37] Xiang Li, Hanzhang Cui, John-Ross Rizzo, Edward Wong, and Yi Fang. 2020. Cross-Safe: A Computer Vision-Based Approach to Make All Intersection-Related Pedestrian Signals Accessible for the Visually Impaired. In *Advances in Computer Vision*, Kohei Arai and Supriya Kapoor (Eds.). Vol. 944. Springer International Publishing, Cham, 132–146. https://doi.org/10.1007/978-3-030-17798-0_13

[38] Yimin Lin, Kai Wang, Wanxin Yi, and Shiguo Lian. 2019. Deep Learning Based Wearable Assistive System for Visually Impaired People. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW).* IEEE, Seoul, Korea (South), 2549–2557. https://doi.org/10.1109/ICCVW.2019.00312

[39] Liza Lin, Newley Purnell. 2019. A World With a Billion Cameras Watching You Is Just Around the Corner. https://www.wsj.com/articles/a-billion-surveillance-cameras-forecast-to-be-watching-within-two-years-11575565402

[40] Sergio Mascetti, Dragan Ahmetovic, Andrea Gerino, and Cristian Bernareggi. 2016. ZebraRecognizer: Pedestrian crossing recognition for people with visual impairment or blindness. *Pattern Recognition* 60 (Dec. 2016), 405–419. https://doi.org/10.1016/j.patcog.2016.05.002

[41] MIPsoft. 2016. BlindSquare. https://www.blindsquare.com/

[42] Marko Modsching, Ronny Kramer, and Klaus ten Hagen. 2006. Field trial on GPS Accuracy in a medium size city: The influence of built-up. In *3rd workshop on positioning, navigation and communication*, Vol. 2006. 209–218.

[43] MQTT. 2022. MQTT: The Standard for IoT Messaging. https://mqtt.org/

[44] Madoka Nakajima and Shinichiro Haruyama. 2012. Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing. In *2012 1st IEEE International Conference on Communications in China (ICCC).* 524–529. https://doi.org/10.1109/ICCChina.2012.6356940

[45] Nvidia. 2023. Nvidia DeepStream GStreamer Plugin: NvDCF Tracker. https://docs.nvidia.com/metropolis/deepstream/dev-guide/text/DS_plugin_gst-nvtracker.html#nvdcf-tracker

[46] NVIDIA. 2023. NVIDIA DeepStream SDK Developer Guide. https://docs.nvidia.com/metropolis/deepstream/dev-guide/index.html

[47] NVivo. 1997. NVivo. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

[48] Petr Oščádal, Daniel Huczala, Jan Bém, Václav Krys, and Zdenko Bobovský. 2020. Smart Building Surveillance System as Shared Sensory System for Localization of AGVs. *Applied Sciences* 10, 23 (Nov. 2020), 8452. https://doi.org/10.3390/app10238452

[49] Sabrina A. Paneels, Dylan Varenne, Jeffrey R. Blum, and Jeremy R. Cooperstock. 2013. The Walking Straight Mobile Application: Helping the Visually Impaired Avoid Veering. (July 2013).

[50] Jagannadh Pariti, Vinita Tibdewal, and Tae Oh. 2020. Intelligent Mobility Cane - Lessons Learned from Evaluation of Obstacle Notification System using a Haptic

Approach. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems.* ACM, Honolulu HI USA, 1–8. https://doi.org/10.1145/3334480.3375217

[51] Roman Pflugfelder and Horst Bischof. 2010. Localization and Trajectory Reconstruction in Surveillance Cameras with Nonoverlapping Views. *IEEE transactions on pattern analysis and machine intelligence* 32 (April 2010), 709–21. https://doi.org/10.1109/TPAMI.2009.56

[52] Giorgio Presti, Dragan Ahmetovic, Mattia Ducci, Cristian Bernareggi, Luca Ludovico, Adriano Baratè, Federico Avanzini, and Sergio Mascetti. 2019. WatchOut: Obstacle Sonification for People with Visual Impairment or Blindness. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility.* ACM, Pittsburgh PA USA, 402–413. https://doi.org/10.1145/3308561.3353779

[53] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. http://arxiv.org/abs/2103.00020 arXiv:2103.00020 [cs].

[54] Lisa Ran, Sumi Helal, and Steve Moore. 2004. Drishti: an integrated indoor/outdoor blind navigation system and service. In *Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. Proceedings of the.* IEEE, 23–30.

[55] Dipankar Raychaudhuri, Ivan Seskar, Gil Zussman, Thanasis Korakis, Dan Kilper, Tingjun Chen, Jakub Kolodziejski, Michael Sherman, Zoran Kostic, Xiaoxiong Gu, et al. 2020. Challenge: COSMOS: A city-scale programmable testbed for experimentation with advanced wireless. In *Proc. ACM MobiCom.*

[56] Manaswi Saha, Alexander J. Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2019. Closing the Gap: Designing for the Last-Few-Meters Wayfinding Problem for People with Visual Impairments. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility.* ACM, Pittsburgh PA USA, 222–235. https://doi.org/10.1145/3308561.3353776

[57] Daisuke Sato, Uran Oh, João Guerreiro, Dragan Ahmetovic, Kakuya Naito, Hironobu Takagi, Kris M. Kitani, and Chieko Asakawa. 2019. NavCog3 in the Wild: Large-scale Blind Indoor Navigation Assistant with Semantic Features. *ACM Transactions on Accessible Computing* 12, 3 (Sept. 2019), 1–30. https://doi.org/10.1145/3340319

[58] Sendero Group. 2019. Seeing Eye GPS. http://www.senderogroup.com/products/seeingeyegps/index.html

[59] Jae Shim and Young Cho. 2016. A Mobile Robot Localization via Indoor Fixed Remote Surveillance Cameras. *Sensors* 16, 2 (Feb. 2016), 195. https://doi.org/10.3390/s16020195

[60] Jae-Hong Shim and Young-Im Cho. [n. d.]. A Mobile Robot Localization using External Surveillance Cameras at Indoor. 56 ([n. d.]), 502–507. https://doi.org/10.1016/j.procs.2015.07.242

[61] Hojun Son, Divya Krishnagiri, V. Swetha Jeganathan, and James Weiland. 2020. Crosswalk Guidance System for the Blind. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC).* 3327–3330. https://doi.org/10.1109/EMBC44109.2020.9176623 ISSN: 2694-0604.

[62] Wim Taymans, Steve Baker, Andy Wingo, Rondald S Bultje, and Stefan Kost. 2013. Gstreamer application development manual (1.2. 3). *Publicado en la Web* 72 (2013).

[63] Juan Terven and Diana Cordova-Esparza. 2023. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. *arXiv preprint arXiv:2304.00501* (2023).

[64] The Royal Institution for the Advancement of Learning (McGill University). 2017. Autour. http://autour.mcgill.ca/en/

[65] F. Van Diggelen and P. Enge. 2015. The world's first GPS MOOC and worldwide laboratory using smartphones, Vol. 1. 361–369.

[66] Charles Vicek, Patricia McLain, and Michael Murphy. 1993. GPS/dead reckoning for vehicle tracking in the" urban canyon" environment. In *Proceedings of VNIS'93-Vehicle Navigation and Information Systems Conference.* IEEE, 461–34.

[67] Hsueh-Cheng Wang, Robert K. Katzschmann, Santani Teng, Brandon Araki, Laura Giarre, and Daniela Rus. 2017. Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. In *2017 IEEE International Conference on Robotics and Automation (ICRA).* IEEE, Singapore, Singapore, 6533–6540. https://doi.org/10.1109/ICRA.2017.7989772

[68] Shiyun Yang, Emily Bailey, Zhengye Yang, Jonatan Ostrometzky, Gil Zussman, Ivan Seskar, and Zoran Kostic. 2020. COSMOS smart intersection: Edge compute and communications for bird's eye object tracking. In *Proc. SmartEdge.*

[69] Chris Yoon, Ryan Louie, Jeremy Ryan, MinhKhang Vu, Hyegi Bang, William Derksen, and Paul Ruvolo. 2019. Leveraging Augmented Reality to Create Apps for People with Visual Disabilities: A Case Study in Indoor Navigation. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility.* ACM, Pittsburgh PA USA, 210–221. https://doi.org/10.1145/3308561.3353788