Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2023.0322000

SLAM for Visually Impaired People: a Survey

Marziyeh Bamdad[1,2], Davide Scaramuzza[1], and Alireza Darvishy[2]
[1]Department of Informatics, University of Zurich, 8050 Zurich, Switzerland
[2]Institute of Applied Information Technology, Zurich University of Applied Sciences, 88400 Winterthur, Switzerland

Corresponding author: Marziyeh Bamdad (e-mail: bamdad@ifi.uzh.ch).

ABSTRACT In recent decades, several assistive technologies have been developed to improve the ability of blind and visually impaired (BVI) individuals to navigate independently and safely. At the same time, simultaneous localization and mapping (SLAM) techniques have become sufficiently robust and efficient to be adopted in developing these assistive technologies. We present the first systematic literature review of 54 recent studies on SLAM-based solutions for blind and visually impaired people, focusing on literature published from 2017 onward. This review explores various localization and mapping techniques employed in this context. We systematically identified and categorized diverse SLAM approaches and analyzed their localization and mapping techniques, sensor types, computing resources, and machine-learning methods. We discuss the advantages and limitations of these techniques for blind and visually impaired navigation. Moreover, we examine the major challenges described across studies, including practical challenges and considerations that affect usability and adoption. Our analysis also evaluates the effectiveness of these SLAM-based solutions in real-world scenarios and user satisfaction, providing insights into their practical impact on BVI mobility. The insights derived from this review identify critical gaps and opportunities for future research activities, particularly in addressing the challenges presented by dynamic and complex environments. We explain how SLAM technology offers the potential to improve the ability of visually impaired individuals to navigate effectively. Finally, we present future opportunities and challenges in this domain.

INDEX TERMS Navigation, SLAM, systematic literature review, visually impaired

I. INTRODUCTION

In recent decades, there has been increasing research interest in developing assistive technologies to enhance spatial navigation for blind and visually impaired (BVI) individuals. In most cases, the main goal is to guide and assist BVI people in navigating safely in unknown environments without the help of a sighted assistant. Navigation is a complex task; it requires finding an optimal path to the desired destination, perceiving the surroundings, and avoiding obstacles. Crucially, all of these functionalities need to accurately localize the BVI user in the environment. There are several approaches for localization, such as the global positioning system (GPS), radio frequency identification (RFID), and simultaneous localization and mapping (SLAM) [1], [2]. Each has advantages and challenges and is used in different applications.

GPS is a localization technique employed in outdoor scenarios owing to its affordability to the end user, wide coverage of the Earth, and ease of integration with other technologies. However, this technique suffers from limitations like satellite signal blockage, inaccuracy, and signal loss caused by weather conditions, walls, and other obstacles [2]. Approaches based on RFID utilize small, low-cost tags for localization. To localize an agent, a set of RFID tags must be installed in the environment [1]. Although localization can be accurately performed using an RFID scheme, taking advantage of this technique requires a pre-installed infrastructure.

A SLAM approach can offer a reliable alternative to RFID and GPS. SLAM is an innovative technique that involves simultaneously constructing an environment model (map) and estimating the state of an agent moving within it [3]. The SLAM architecture (Figure 1) consists of two fundamental components: the front-end and the back-end. The front-end receives environmental information from the sensors, abstracts it into amenable models for estimation, and sends it to the back-end [3]. The back-end is responsible for optimizing the mapping, localization, and data fusion processes, which collectively contribute to the accuracy and reliability of the SLAM systems.

SLAM uses diverse types of sensors to determine an agent's position, location, and velocity and detect and avoid obstacles, even in a dynamically changing unknown environment. This technique uses infrared (IR) sensors, acoustic

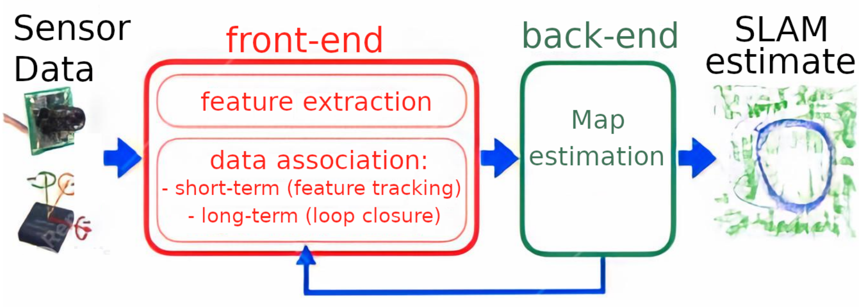

FIGURE 1. Front- and back-end in a typical SLAM system [3]

sensors, RGB cameras, inertial measurement units (IMUs), ultrawide-band (UWB), LiDAR, RADAR, and RGB-D sensors [4].

The collaborative effort between the front- and back-ends empowers SLAM to provide a robust and real-time spatial understanding, making it a valuable tool for various applications. The SLAM community has made tremendous strides over the past 35 years, developing large-scale practical applications and seeing a steady transition of this technology into the industry [3].

SLAM technology has been widely applied in various fields, demonstrating its versatility and robustness. In autonomous driving, SLAM enables vehicles to create and update maps of their environments in real time [130]. This capability is essential for safe and efficient road navigation, allowing accurate localization and obstacle avoidance in dynamic urban environments. In robotics, SLAM is crucial for navigation [131]. This allows robots to operate effectively in unknown environments by simultaneously building a map of their surroundings and determining their positions within them. Such a capability is essential for autonomous exploration, path planning, and obstacle avoidance in diverse settings, ranging from indoor spaces to outdoor terrains. Augmented reality (AR) relies heavily on SLAM for the accurate placement of virtual objects in the real world, enhancing user experiences in gaming, education, and industrial applications [132]. In underwater robotics, SLAM helps autonomous underwater vehicles navigate highly unstructured and complex marine environments and supports exploration, research, and maintenance [133]. For aerial vehicles, SLAM is an indispensable methodology for autonomous flight performed by unmanned aerial vehicles (UAVs) [134], along with flight control [135], [136]. SLAM enables drones to navigate and map areas autonomously, which is valuable for tasks such as search-and-rescue, surveying, and environmental monitoring. These diverse applications highlight the versatility of SLAM and its critical role in enabling autonomous operation across various scenarios and platforms, from urban landscapes to ocean depths.

The evolution of portable computation and the availability of low-cost, highly accurate, and lightweight sensors such as cameras and IMUs make them appropriate for pedestrian navigation. By exploiting these advances, many researchers have recently adopted SLAM to develop assistive technology demonstrators to help BVI people navigate unknown environments.

Since the first electronic travel aids (ETAs) emerged approximately 70 years ago, the development of navigation devices to guide BVI people through indoor and/or outdoor environments has remained a challenge and a key concern for researchers [5]. From traditional to deep-learning-based navigation approaches, researchers have always faced challenges ranging from technical issues to the limitations of user capabilities. As BVI navigation approaches must improve real-time performance while reducing the size, weight, energy cost, and overall price of the assistive system, these studies have put a lot of effort into coping with constraints in computational issues, sensory equipment, and portable devices. They also need to provide solutions to calculate the precise position and orientation of the user in a real-time manner. However, the challenges of different scenarios, including complex and cluttered environments, noisy environments, and large spaces, must be considered.

Furthermore, efficient and reliable obstacle detection in both indoor and outdoor environments has always been a concern. In this regard, other challenges include identifying static and dynamic obstacles, predicting the risk of collision, understanding moving objects' motion and estimating their speed, detecting small objects, and identifying obstacles at different levels of the user's body, from drops in terrain to head level. In addition, an intuitive, user-friendly, low-cognitive-load method to provide accurate and sufficient environmental information to the user is also considered an important research target. These methods should be improved to provide adjustable and customized feedback on demand for different users.

Moreover, assistive technology should provide user safety and independence, hands-free operations, decreased effort, and backup in the case of system failure. In addition to the aforementioned challenges, deep-learning-based solutions also have special issues, such as designing lightweight neural network architectures to reduce computational expense and provide sufficient data for the training and validation of the models.

This SLR is designed to act as a resource for the academic and research communities. The objective of this review is to explore and highlight the strengths and potential limitations of the current SLAM applications for visually impaired navigation. This study aims to inform and guide subsequent research. The insights derived from this review identify critical gaps and opportunities for future research, particularly for tackling the challenges presented by dynamic and complex environments. Such environments pose unique difficulties for visually impaired navigation, and addressing them through advanced SLAM technologies could lead to significant improvements in both the effectiveness and reliability of assistive solutions.

A. RELATED WORK

Thus far, many reviews have been conducted on assistive technologies developed for BVI navigation. Several studies reviewed walking assistance systems [5], [7]–[14] provided a detailed classification of the developed approaches. [12] categorized walking assistants into three groups: sensor-based, computer vision-based, and smartphone-based. The authors explained the technologies used and inspected each approach, and evaluated some important parameters of each approach, such as the type of capturing device, type of feedback, working area, cost, and weight. The work by [14] introduced techniques and technologies designed to assist visually impaired individuals in their mobility and daily lives. This comprehensive review analyzes multiple mobility-assistive technologies that are suitable for indoor and outdoor environments. It offers insights into the various feedback methods employed by assistive tools based on recent technologies. In addition, [7] reviewed wayfinding devices used by visually impaired individuals in real-world scenarios. This review aimed to provide a comprehensive exploration of the various aids employed for navigation while assessing their perceived efficacy.

Some studies focused on indoor navigation for BVI users [15]–[22] and some focused on computer vision-based navigation systems [15], [23]–[27]. Among these studies, [15] conducted a systematic literature review of state-of-the-art computer vision-based methods used for indoor navigation. The authors described the advantages and limitations of each solution under review, and included a brief description of each method. Furthermore, [21] comprehensively examined existing methods and systems developed within the domain of assistive technology, with a specific focus on addressing the unique needs and challenges faced by the visually impaired. This study places strong emphasis on evaluating methods that have practical applications in enhancing the lives of visually impaired individuals.

Several review papers on wearable navigation systems have also been published [28]–[33]. [28] have conducted a systematic review with the primary objectives of analyzing wearable obstacle avoidance electronic travel aids. Their work delves into the strengths and weaknesses of existing ETAs, providing a thorough evaluation of hardware functionality, cost-effectiveness, and the overall user experience. [29] provided a comprehensive understanding of wearable travel aids by focusing on their designs and usability. Their objectives included surveying the current landscape of travel aid design, investigating key design issues, and identifying limitations and future research directions. Furthermore, [30] conducted a systematic review of the literature on wearable technologies designed to enhance the orientation and mobility of the visually impaired. This review provides valuable insights into the technological characteristics of wearables, identifies feedback interfaces, emphasizes the importance of involving visually impaired individuals in prototype evaluations, and highlights the critical need for safety evaluations. A review by [31] provides a comprehensive review of computer vision and machine-learning-based assistive methods. Existing ETAs are divided into two groups: active systems providing subject localization and object identification, and passive systems providing information about the users' surroundings using a stereo camera, monocular camera, or RGB-D camera.

Focusing on guide robots, [34] reviewed the multifaceted objectives. Their work included a comparative analysis of the existing robotic mobility aids and state-of-the-art technologies. This review highlights the potential of guide robots to enhance the mobility and independence of the visually impaired.

[35] and [36] reviewed studies with the focus on object detection and recognition. [36] performed a review on object recognition tailored to the needs of visually impaired individuals. This review examines state-of-the-art object detection and recognition techniques, focuses on standard datasets, and emphasizes on the latest advancements. [35] reviewed studies specific for staircase detection systems, primarily designed to facilitate the navigation of visually impaired individuals. The goal of this review is to provide a comprehensive comparative analysis of these systems considering their suitability and effectiveness.

Other similar studies include a survey of inertial measurement units (IMUs) in assistive technologies for visually impaired people [37], a review of urban navigation for BVI people [38], a survey paper that reviewed assistive tools based on white canes [39], and review papers exploring smartphone-based navigation devices [40]–[42]. [40] reviewed the multifaceted objectives in the domain of smartphone-based navigation devices. They aimed to provide a comprehensive overview of smartphone use among people with vision impairment, identify research gaps for future exploration, and delve into the use of smartphones by individuals with vision impairment and the accessibility challenges they encounter. To the best of our knowledge, there is no survey paper on SLAM-based navigation systems for BVI people. Our study aims to bridge this gap in literature.

B. CONTRIBUTION

This paper presents a systematic literature review (SLR) addressing fundamental questions regarding SLAM-based approaches for BVI navigation. This review provides insights into technological diversity, advantages, limitations, and the potential to address real-world challenges. While recognizing the broad range of potential research questions we narrowed our focus to the four questions outlined in Table 2. The primary contributions of this study are as follows:

- Identification of SLAM approaches: We systematically identified and categorized the diverse SLAM approaches adopted in the development of assistive systems tailored for visually impaired navigation. This includes analyzing the localization and mapping techniques, sensor types, computing resources, and machine-learning methods used in these approaches.
- Advantages and limitations synthesis: Our study synthesizes the advantages and limitations of these SLAM techniques when applied to BVI navigation.
- Classification of challenges: We identify and categorize studies that address challenging conditions relevant to SLAM-based navigation systems for the visually impaired. In addition, we discuss practical considerations that affect the usability and adoption of these systems.
- Exploration of the potential for enhancing BVI navigation: We analyzed how the proposed SLAM-based approaches improved navigation in visually impaired individuals. In addition, we evaluated the effectiveness of these solutions in real-world scenarios and assessed user satisfaction to understand their practical impact on BVI mobility.

C. PAPER STRUCTURE

The remainder of this paper is organized as follows. In Section II, we explain the protocol, methodology, tools, and techniques used to conduct SLR. The findings of our SLR and answers to the SLR research questions are summarized in Section III. Section IV presents the future opportunities and potential advancements in this domain. Finally, Section V concludes the paper.

II. SLR METHODOLOGY

A systematic literature review is one of the most common types of literature review used to collect, review, appraise, and report research studies on a specific topic, adhering to predefined rules for conducting the review [43]. Compared with traditional literature reviews, it provides a wider and more precise understanding of the topic under review [44]. Various guidelines exist for conducting SLR in different research fields such as software engineering [45]–[47], computer science [48], information systems [49], planning education and research [50], and health sciences [44], [51]. To conduct this review, we followed the guidelines for conducting systematic reviews proposed by [52]. Figure 2 illustrates our SLR process.

SLR consists of three key phases: planning, conducting, and reporting the review. We defined our research questions and motivation, keywords, and search string, as well as selection criteria in the planning phase of the SLR. In the conducting phase, we executed searches on digital sources using predefined search strings that were established during the planning stage. We evaluated the quality of the selected papers and extracted relevant data aligned with SLR research questions.

We used the PICOC criteria to identify the key elements that needed to be considered and frame our research questions. PICOC represents Population, Intervention, Comparison, Outcome, and Context [53]. Table 1 lists the PICOC elements, relevant values, and descriptions of these elements in our study.

In this section, we first introduce the tool we used to manage our SLR process and then detail our methodology for conducting our systematic literature review in planning and conducting the review subsections.

A. SLR TOOL

Various tools have been used to conduct systematic literature reviews. Some of them are commercial such as Covidence[1], DistillerSR[2], and EPPI-Reviewer[3]; and some are free such as Cadima[4], Rayyan[5], RevMan[6], and Parsifal[7]. We used the Parsifal platform to manage the SLR phases. It is an online tool developed to support the process of performing SLR. Parsifal provides researchers with an interface to invite co-authors to collaborate in a shared workspace on the SLR. During the planning phase, this tool assists the authors by addressing the objectives, PICOC, research questions, search strings, keywords and synonyms, selection of sources, and inclusion and exclusion criteria. Parsifal offers tools for creating a quality assessment checklist and data extraction forms. In the conducting phase, this tool helps the authors import the bibtex files and select studies. It assists in identifying and eliminating duplicates among various sources, performing quality assessments, and facilitating data extraction from papers. Finally, it provides a method to document the entire SLR process.

B. PLANNING THE REVIEW

The first step in conducting SLR is to establish a protocol. The protocol outlines the review procedures and ensures replicability. Within the protocol, we formulated our research questions, designed a search strategy, and defined the specific criteria for selecting relevant studies. In addition, we defined a set of criteria presented in Table 8 to evaluate the quality of the selected literature. Furthermore, to facilitate the extraction of data in alignment with our research questions, we designed a data-extraction form.

[1] https://www.covidence.org/home

[2] https://www.evidencepartners.com/products/distillersr-systematic-review-software/

[3] https://eppi.ioe.ac.uk/CMS/Default.aspx

[4] https://www.cadima.info/

[5] https://rayyan.qcri.org/welcome

[6] https://training.cochrane.org/online-learning/core-software-cochrane-reviews/revman

[7] https://parsif.al

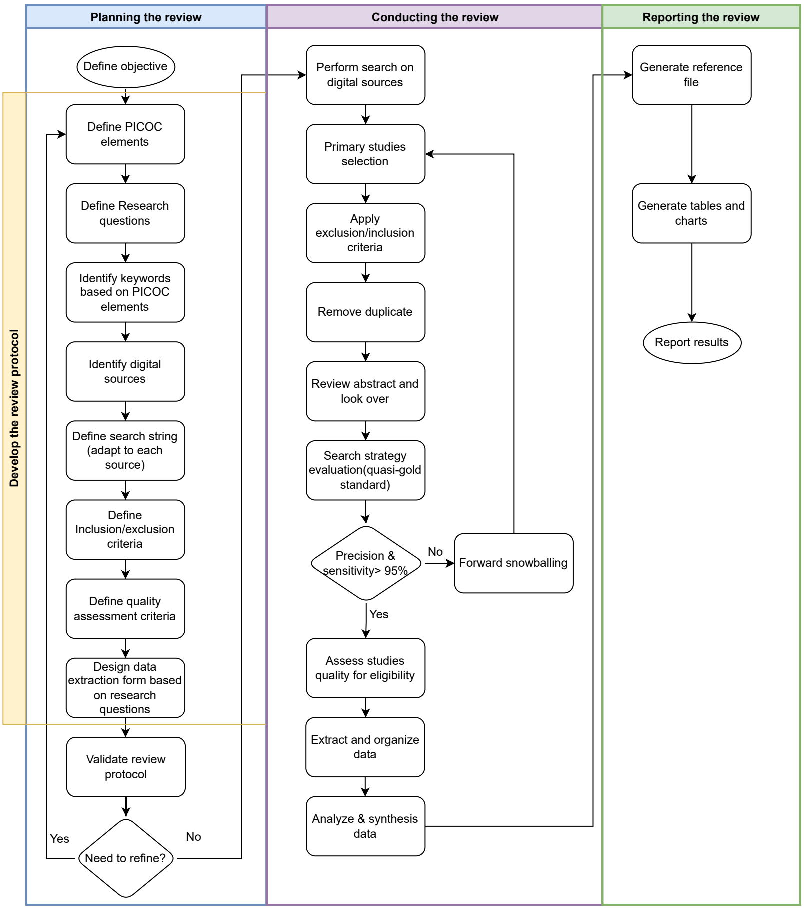

FIGURE 2. The process of SLR

TABLE 1. Elements of PICOC

Element	Description [53]	Value
Population	The problem or situation	Visually impaired navigation
Intervention	The technology, tool or method under study	SLAM
Comparison (optional)	The technology, tool or method with which the intervention is compared	-
Outcome	Results that Intervention could produce	Lightweight, affordable, accurate, efficient assistive technology
Context	The specific context of the study	Autonomous mobility of visually impaired people

1) Research Questions and Motivation

The SLAM technique is widely used for the navigation of robots, autonomous drones, and self-driving cars, owing to its performance, reliability, and efficiency. Therefore, we reviewed the literature on visually impaired navigation, which was designed based on the SLAM technique. Our aim was to determine the advantages and limitations of employing this technique for visually impaired navigation as well as to identify opportunities for future research. Furthermore, we aimed to explore how extensively this method has been used in this specific area of research. Table 2 presents the research questions that guided this review, and a description of the questions.

2) Search strategy

A key step in performing SLR is to design an effective search strategy. This strategy should be executed with reasonable effort to retrieve relevant studies from digital libraries [54]. The exhaustive search process in systematic reviews is a critical factor distinguishing them from traditional literature reviews [54], leading to a wider and more precise understanding of the topic under review. To design the search string we first extracted keywords from the PICOC elements, including population, intervention, and outcomes. We then determined synonyms for each keyword to broaden the search string. The list of keywords and their synonyms related to each PICOC element is listed in Table 3.

The first part of the search string (i.e. 'visual* impair*' OR 'blind' OR 'visually disabled' OR 'sight impaired') is

TABLE 2. Research questions for the SLR process

No.	Research question	Description
RQ1	What localization and mapping approaches are used for the navigation of the visually impaired?	The target is to identify different localization and mapping techniques adopted for the development of assistive devices for visually impaired navigation.
RQ2	What are the advantages and limitations of SLAM techniques for BVI navigation?	The objective is to summarize the advantages and constraints of SLAM-based approaches for visually impaired navigation.
RQ3	What challenging situations have been addressed?	The purpose of this question is to know which of the challenging conditions (e.g. crowded environment, changing view point, challenging light conditions, dynamic objects, etc.) relevant to navigation systems have been considered.
RQ4	How does the proposed solution improve navigation using SLAM for individuals with impaired vision?	The research question seeks to understand how SLAM techniques can enhance mobility and navigation in individuals with visual impairments.

TABLE 3. Keywords used to design search string

Keywords	Synonyms	PICOC element
Visually impaired navigation	Blind navigation, Navigation assistance for the visually impaired, Navigational guidance for individuals with visual impairments, Navigational support for the visually disabled, Orientation and mobility for the visually impaired, Sight-impaired navigation, Visual impairment navigation aid, low vision navigation, partially sighted navigation	Population
SLAM	Real-time mapping and positioning, Simultaneous Localization and Mapping, Simultaneous mapping and position tracking, mapping and localization	Intervention
Accurate, efficient, reliable assistive technology	high-performance, precise, effective, trustworthy assistive technology	Outcome

relevant to the population element of the PICOC framework. The addition of an asterisk to the terms 'visual' and 'impair' allows us to include various expressions, including 'visually impaired' and 'visual impairment.' The following segment of the query, consisting of 'navigation*' OR 'mobility' OR 'wayfinding,' also places emphasis on the population aspect within the context of the systematic review. The inclusion of an asterisk in 'navigation*' ensures comprehensive coverage, accounting for variations such as 'navigational.' Regarding the Intervention component defined in the PICOC framework, we employed the term 'SLAM' in conjunction with synonyms identified in the literature from diverse domains where SLAM is applied, such as robotics, autonomous driving cars, and underwater SLAM. The last segment of the search string is connected to the outcome element of the PICOC. Adding keywords such as "localization" alone or the specific names of SLAM techniques did not increase the number of related papers. The search strings were employed on ten large citation databases, as listed in Table 4, to carry out an exhaustive search. We modified the base search string according to the Search Tip in each library to satisfy specific requirements.

We utilized the Advanced Search feature in digital libraries to gain more control over our search parameters. The title, abstract, and keyword fields were selected to retrieve the search results. Searching on Google Scholar is somewhat different from searching for other digital libraries. Unlike other platforms, Google Scholar does not suggest various filters, requiring the manual incorporation of filters into the search string. Additionally, to identify English-written studies, we adjusted the language preference settings within our Google Scholar account to filter the search results in English.

3) Selection criteria

Table 5 presents the selection criteria used to identify the eligible studies during the selection process. The Availability criterion included studies accessible in full text from digital databases. In addition, the Language criterion ensured the inclusion of publications written only in English. Furthermore, the Publication Period criterion restricted the inclusion of studies published between January 2017 and July 2023. This timeframe was chosen to prioritize the most current and state-of-the-art approaches in this rapidly evolving field. By focusing on this recent period, we aimed to provide a comprehensive yet manageable review of the latest innovations without overwhelming readers with potentially outdated information. The Type of Source criterion included conference and journal papers, which were considered peer-reviewed and academically recognized sources. Books, dissertations, newsletters, speeches, technical reports, and white papers were excluded. Finally, the Relevance criterion played an important role in the exclusion process; therefore, publications outside the scope of our study were excluded based on a review of their titles and abstracts.

C. CONDUCTING THE REVIEW

As shown in Figure 2, the review process began after the review protocol was finalized. The conducting phase is a multi-stage process that includes research identification, study selection, data extraction, and data synthesis. In the research identification step, digital libraries were searched using adapted search strings that were specific to each library. This search aimed to collect a pool of potentially relevant primary studies. The next step involved the selection of studies for which the relevance of each study to the review was evaluated. The steps involved in this process are illustrated in

TABLE 4. Databases selected for the search procedure

Digital source	Web address	# of papers	Last access date
ACM Digital Library	http://portal.acm.org	4	22 Jul 2023
Google Scholar	https://scholar.google.com	0	23 Jul 2023
IEEE Xplore	http://ieeexplore.ieee.org	11	22 Jul 2023
MDPI	https://www.mdpi.com	2	23 Jul 2023
PubMed	https://www.ncbi.nlm.nih.gov/pubmed	3	22 Jul 2023
Science Direct	http://www.sciencedirect.com	2	22 Jul 2023
Scopus	http://www.scopus.com	22	23 Jul 2023
Springer Link	http://link.springer.com	2	22 Jul 2023
Taley & Francis	https://www.tandfonline.com	0	22 Jul 2023
Wiley Online Library	https://onlinelibrary.wiley.com	1	22 Jul 2023

TABLE 5. Selection criteria

Criteria	Inclusion	Exclusion
Availability	Available in full text	Not accessible in specific databases
Language	English	Not written in English
Publication period	From 2017 to July 2023	Prior 2017
Type of source	Conferences and journals papers	Books, dissertation, newsletters, speeches, technical reports, white papers
Relevance	Papers relevant to at least two research questions	Outside the scope of our research

Figure. 3. During the data extraction phase, the data required from the studies were collected and analyzed. We employed the data extraction form established during the development phase of the review protocol to ensure accurate extraction of information that addresses our research questions.

1) Identification

During the initial phase of our review, we conducted searches across the digital libraries using custom-formulated queries for each library. For each dataset, we ran three different search strings (SS1, SS2, and SS3), as shown in Table 6, consisting of various combinations of keywords, booleans, and wildcard operators. These search strings were applied to all digital libraries except Google Scholar. For Google Scholar, we initially used keywords similar to those used in SS1, resulting in over 11,000 results. Upon reviewing a subset of these, we determined that a significant number were not relevant to our topic. Consequently, we decided to use only the primary keywords (shown in Table 3) to construct the search string for this digital library.

We selected the search string that yielded the most results to identify primary studies and then applied exclusion criteria to the results obtained from the search strings used for each library. We observed that SS3, which incorporated 'Orientation and mobility' to refine the search by focusing on more specialized literature, did not yield better results than SS1, which included the general term 'mobility', across all digital libraries. This indicates that the broader term 'mobility' was sufficient to capture the necessary literature. The specificity of SS3 did not contribute to additional relevant results. Additionally, upon receiving the message 'Use fewer Boolean connectors (maximum 8 per field)' while running SS1 on ScienceDirect, we switched to SS2 to maintain the number of Boolean connectors within the limit.

The initial searches of all digital libraries resulted in 6,809 records. The search strings used for each digital library is presented in Table 7.

The results obtained from digital libraries searches were exported in the BibTex format, a process facilitated by the export citation features available in the libraries. The BibTex data were then imported into the Parsifal framework for subsequent stages of our review. Springer and Google Scholar do not provide direct options for exporting data in BibTex format. To address this issue, we used Zotero and its browser plugin, Zotero Connector, to streamline the process. With these tools, we added paper information from webpage views to Zotero and subsequently retrieved BibTex data.

For Springer Link, which provides only CSV files with search results, we opened the CSV in Excel and extracted the DOIs. These DOIs were then pasted into Zotero's "Add item(s) by identifier" feature. After importing the DOIs into Zotero, we selected the appropriate folder containing the imported papers and exported the collection to the BibTex format using a simple right-click. As Scholar Google does not provide easy export of a large number of records, we adopted a similar approach: creating a library, saving search results to that library, and exporting paper data in BibTex format from that library. This process ensured that we obtained the necessary data for the subsequent stages of our systematic review.

2) Study selection

After conducting searches in the digital libraries, we applied our selection criteria, as defined in our review protocol, to filter out irrelevant studies. Initially, records published before 2017 were excluded. Further exclusions involved filtering out publications that were not written in English or had not been published in peer-reviewed venues. Following these steps, of the initial 6809 records found in the initial search, 5431 were excluded.

We imported the study data into the Parsifal platform in BibTeX format, as explained in Section II-C1, which helped

TABLE 6. Search strings applied to digital libraries, featuring keywords and operators to identify primary studies.

SS1	SS2	SS3	SS4
("Visually impaired" OR blind OR "visual impairment*" OR "visually disabled" OR "Sight impaired") AND (navigation* OR mobility) AND (SLAM OR "Simultaneous Localization and Mapping" OR "Real-time mapping and positioning" OR "Simultaneous mapping and position tracking" OR "mapping and localization") AND (technology OR aid OR support OR assist*)	("Visually impaired" OR blind) AND (navigation OR mobility) AND (SLAM OR "Simultaneous Localization and Mapping")	("Visually impaired" OR blind OR "visual impairment*" OR "visually disabled" OR "Sight Impaired") AND (navigation* OR **"Orientation and mobility"**) AND (SLAM OR "Simultaneous Localization and Mapping" OR "Real-time mapping and positioning" OR "Simultaneous mapping and position tracking" OR "mapping and localization") AND (technology OR aid OR support OR assist*)	("conference paper" OR "journal") AND ("Visually impaired") AND (navigation) AND (SLAM) AND (-review) AND (-survey)

TABLE 7. Utilization of search strings by digital libraries

Digital source	Search string	Number of results
ACM Digital Library	SS1	284
Google Scholar	SS4	602
IEEE Xplore	SS2	50
MDPI	SS2	191
PubMed	SS1	8
Science Direct	SS2	518
Scopus	SS2	1510
Springer Link	SS2	2585
Taley & Francis	SS2	486
Wiley Online Library	SS2	575

TABLE 8. Quality assessment criteria and weights

Criteria	Weight
Is there an adequate description of the context in which the research was carried out?	0.0, 0.5, 1.0
Does the methodology take into account both localization and mapping issues?	0.0, 0.1, 2.0
Is the solution proposed well presented?	0.0, 0.5, 1.0
Is there a clear statement of findings?	0.0, 0.5, 1.0
Is the research design appropriate to address the aims of the research?	0.0, 0.5, 1.0
Does the study add value to the research community?	0.0, 0.5, 1.0

remove duplicate studies. A total of 116 duplicate papers were excluded. We then reviewed the titles and abstracts of the remaining studies, excluding those irrelevant to our research topic. In this step, 779 studies were excluded.

In the next step, we performed a fast reading of the full text of the remaining papers, excluding 265 studies that were outside the scope of our research. We then evaluated the quality of the studies based on the quality assessment criteria defined in the SLR protocol. Five studies were removed during the assessment of study quality. Table 8 lists the quality assessment criteria for our SLR.

We carefully read 213 full-text articles to address the research questions. As 166 articles were not relevant to at least two of our research questions, they were removed, leaving 47 articles for the final stage.

To objectively assess the performance of our search strategy, we employed the quasi-gold standard (QGS) technique, as described by [54]. Using this method, a set of articles related to the research topic is manually selected. Digital libraries are then searched based on the research strategy to identify related studies. Finally, the retrieved articles are compared with QGS, and the sensitivity of the search strategy is calculated using the following formula:

$$\text{Sensitivity} = \frac{\text{Number of relevant studies retrieved}}{\text{Total number of relevant studies}} \times 100$$

In our SLR, with 30 manually selected relevant studies and 48 studies retrieved using the SLR search strategy, of which 26 were among the manually selected studies, the resulting sensitivity was approximately 86.67%.

To provide a broader range of relevant studies, we included papers on the forward snowballing process [55]. This process involves identifying and accessing references in a paper and reviewing cited papers. We used "Cited by" feature of Google Scholar to identify these additional papers. In this stage, 695 articles were identified. After removing duplicates and applying selection criteria similar to those used for the articles obtained from digital libraries, we added seven more articles to the final collection. Consequently, 54 articles were included in this review. Figure 3 shows a diagram of the study-selection process.

It is important to note that the last search conducted in digital libraries was on July 23, 2023, and that for forward snowballing was on August 12, 2023. These dates should be considered as the starting points for future reviews. The publications included in our review are listed in Tables 9–11 and categorized based on their publication venues. These tables provide an overview of the literature, including the paper title, author names, publication year, location, and source through which the studies were discovered. Among the 54 studies included in our SLR, 27 were sourced from journals, as presented in Table 9. The remaining 27 studies were presented at conferences, as shown in Tables 10 and 11.

Additionally, Tables 12–15 summarize the perspectives and innovations presented in the publications with insights into their limitations and advantages. These tables demonstrate the research issues addressed and the contributions of each study, highlighting the strengths and potential drawbacks of the proposed solutions. They also indicate which solutions are open source, with only seven papers having some or all parts of the project available as open source. Links

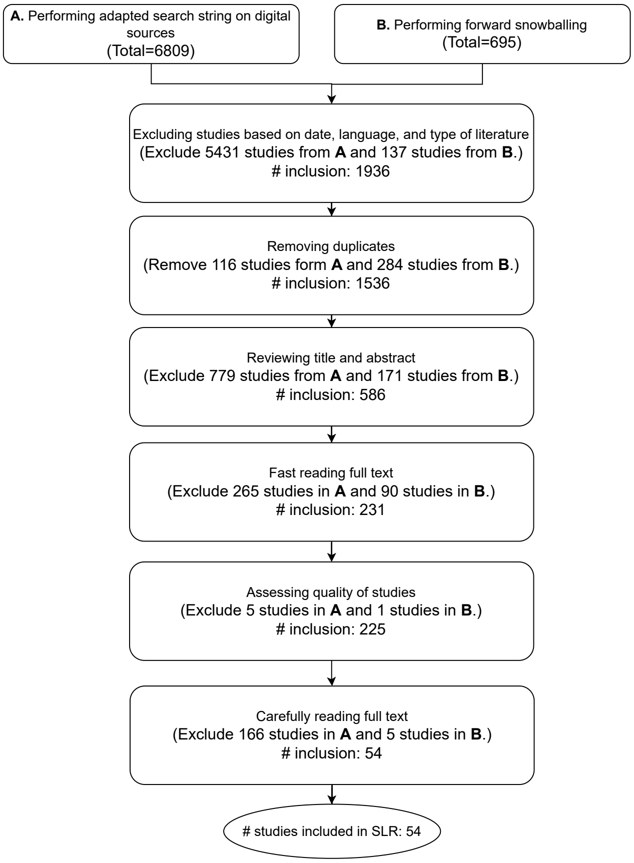

FIGURE 3. Studies selection process

to the sources are provided in these tables if they are directly available in the relevant papers.

3) Data extraction

Data extraction is a critical phase in the systematic literature review process in which relevant data from selected studies are systematically collected. To achieve this objective, we employed the data-extraction form defined in the SLR protocol. This form consists of various fields designed to retrieve answers to our research questions from each of the included articles. Within the scope of this SLR, we defined the following essential elements, each contributing to a comprehensive understanding of the reviewed literature:

- Short summary of the paper: A concise overview of the main points and findings of the study.
- Research issue and contribution: Summary of the research issues addressed and contributions of the studies.
- Localization and mapping technique: Identification of specific techniques applied for localization and mapping.
- Localization and mapping accuracy and robustness: Assessment of accuracy and robustness levels in localization and mapping techniques.
- Running time: Analysis of the running time for localization and mapping techniques.
- Advantages of the presented method: The strengths associated with the localization method presented in each paper for visually impaired navigation.
- Limitations of the presented method: Identification of weaknesses or constraints associated with the localization technique.
- Types of obstacles addressed: Categorization of obstacles, static and dynamic, as a challenge during navigation.
- Challenging conditions: Explanation of other challenging scenarios that the methods are designed to handle.
- Types of sensors: Identification of sensors employed to receive data from the surroundings.
- Computing resources: Identification of computing resources used in SLAM-based solutions.
- Improvement in navigation: Identification of how SLAM-based methods enhance navigation for individuals with impaired vision.
- Working area: Whether the method is intended for indoor, outdoor, or both indoor and outdoor environments.
- Practical challenges and operational efficiency: Evaluation of the user-friendliness, cost-efficiency, weight, comfort for extended use, adjustable fit, fatigue mitigation, and portability of the SLAM-based assistive tools.
- System prototype information: Detailed information on functionalities, sensors, computing resources, human-computer interaction (HCI) mechanisms, assistive tools, and battery life.
- User evaluation: Assessment of user satisfaction of the SLAM-based assistive tools.
- Machine learning techniques: Identification of machine learning techniques used in assistive solutions.
- Open-source availability: Identification of open-source contributions in the reviewed studies.
- Possible future opportunities and directions: Exploration of potential future research areas and directions stemming from these findings.
- The research questions addressed: Identification of specific SLR research questions addressed by each study.

TABLE 9. List of journal publications included in SLR

Ref.	Title	Authors	Published	Year	Source
[56]	Sonification of navigation instructions for people with visual impairment	Dragan Ahmetovic and Federico Avanzini and Adriano Baratè and Cristian Bernareggi and Marco Ciardullo and Gabriele Galimberti and Luca A. Ludovico and Sergio Mascetti and Giorgio Presti	International Journal of Human-Computer Studies	2023	Science@Direct
[57]	Sensing and Navigation of Wearable Assistance Cognitive Systems for the Visually Impaired	Li, Guoxin and Xu, Jiaqi and Li, Zhijun and Chen, Chao and Kan, Zhen	IEEE Transactions on Cognitive and Developmental Systems	2023	IEEE Digital Library
[58]	Mixture reality-based assistive system for visually impaired people	Jucheng Song and Jixu Wang and Shuliang Zhu and Haidong Hu and Mingliang Zhai and Jiucheng Xie and Hao Gao	Displays	2023	Science@Direct
[59]	Research on Design and Motion Control of a Considerate Guide Mobile Robot for Visually Impaired People	Zhang, Bin and Okutsu, Mikiya and Ochiai, Rin and Tayama, Megumi and Lim, Hun-Ok	IEEE Access	2023	Scopus
[60]	UNav: An Infrastructure-Independent Vision-Based Navigation System for People with Blindness and Low Vision	Yang, Anbang and Beheshti, Mahya and Hudson, Todd E and Vedanthan, Rajesh and Riewpaiboon, Wachara and Mongkolwat, Pattanasak and Feng, Chen and Rizzo, John-Ross	Sensors	2022	PubMed
[61]	Multi-Floor Indoor Localization Based on Multi-Modal Sensors	Zhou, Guangbing and Xu, Shugong and Zhang, Shunqing and Wang, Yu and Xiang, Chenlu	Sensors	2022	Scopus
[62]	Knowledge driven indoor object-goal navigation aid for visually impaired people	Hou, Xuan and Zhao, Huailin and Wang, Chenxu and Liu, Huaping	Cognitive Computation and Systems	2022	Wiley Online Library
[63]	Indoor-Guided Navigation for People Who Are Blind: Crowdsourcing for Route Mapping and Assistance	Plikynas, Darius and Indriulionis, Audrius and Laukaitis, Algirdas and Sakalauskas, Leonidas	Applied Sciences (Switzerland)	2022	Scopus
[64]	A Multi-Sensory Guidance System for the Visually Impaired Using YOLO and ORB-SLAM	Xie, Zaipeng and Li, Zhaobin and Zhang, Yida and Zhang, Jianan and Liu, Fangming and Chen, Wei	Information	2022	MDPI
[65]	Egocentric Human Trajectory Forecasting with a Wearable Camera and Multi-Modal Fusion	Qiu, Jianing and Chen, Lipeng and Gu, Xiao and Lo, Frank P.-W. and Tsai, Ya-Yen and Sun, Jiankai and Liu, Jiaqi and Lo, Benny	IEEE Robotics and Automation Letters	2022	Scopus
[66]	A wearable navigation device for visually impaired people based on the real-time semantic visual slam system	Chen, Zhuo and Liu, Xiaoming and Kojima, Masaru and Huang, Qiang and Arai, Tatsuo	Sensors	2021	PubMed
[67]	Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision	Slade, Patrick and Tambe, Arjun and Kochenderfer, Mykel J.	Science Robotics	2021	Scopus
[68]	Assistive Navigation Using Deep Reinforcement Learning Guiding Robot With UWB/Voice Beacons and Semantic Feedbacks for Blind and Visually Impaired People	Lu, Chen-Lung and Liu, Zi-Yan and Huang, Jui-Te and Huang, Ching-I and Wang, Bo-Hui and Chen, Yi and Wu, Nien-Hsin and Wang, Hsueh-Cheng and Giarré, Laura and Kuo, Pei-Yi	Frontiers in Robotics and AI	2021	Scopus
[69]	Indoor Wearable Navigation System Using 2D SLAM Based on RGB-D Camera for Visually Impaired People	Hakim, Heba and Fadhil, Ali	Advances in Intelligent Systems and Computing	2021	Scopus
[70]	An RGB-D Camera Based Visual Positioning System for Assistive Navigation by a Robotic Navigation Aid	Zhang, He and Jin, Lingqiu and Ye, Cang	IEEE/CAA Journal of Automatica Sinica	2021	IEEE Digital Library
[71]	Hierarchical visual localization for visually impaired people using multimodal images	Cheng, Ruiqi and Hu, Weijian and Chen, Hao and Fang, Yicheng and Wang, Kaiwei and Xu, Zhijie and Bai, Jian	Expert Systems with Applications	2021	Scopus
[72]	Indoor Topological Localization Based on a Novel Deep Learning Technique	Liu, Qiang and Li, Ruihao and Hu, Huosheng and Gu, Dongbing	Cognitive Computation	2020	Scopus
[73]	Combining Obstacle Avoidance and Visual Simultaneous Localization and Mapping for Indoor Navigation	Jin, SongGuo and Ahmed, Minhaz Uddin and Kim, Jin Woo and Kim, Yeong Hyeon and Rhee, Phill Kyu	Symmetry	2020	MDPI
[74]	Wearable travel aid for environment perception and navigation of visually impaired people	Bai, Jinqiang and Liu, Zhaoxiang and Lin, Yimin and Li, Ye and Lian, Shiguo and Liu, Dijun	Electronics (Switzerland)	2019	Scopus
[75]	An ARCore based user centric assistive navigation system for visually impaired people	Zhang, Xiaochen and Yao, Xiaoyu and Zhu, Yi and Hu, Fei	Applied Sciences (Switzerland)	2019	Scopus
[76]	Virtual-Blind-Road Following-Based Wearable Navigation Device for Blind People	Bai, Jinqiang and Lian, Shiguo and Liu, Zhaoxiang and Wang, Kai and Liu, Dijun	IEEE Transactions on Consumer Electronics	2018	IEEE Digital Library
[77]	An indoor wayfinding system based on geometric features aided graph SLAM for the visually impaired	Zhang, He and Ye, Cang	IEEE Transactions on Neural Systems and Rehabilitation Engineering	2017	PubMed
[78]	Plane-Aided Visual-Inertial Odometry for 6-DOF Pose Estimation of a Robotic Navigation Aid	Zhang, He and Ye, Cang	IEEE Access	2020	Scopus
[79]	SRAVIP: Smart Robot Assistant for Visually Impaired Persons	Albogamy, Fahad and Alotaibi, Turk and Alhawdan, Ghalib and Mohammed, Faisal	International Journal of Advanced Computer Science and Applications	2021	Forward Snowballing
[80]	A Lightweight Approach to Localization for Blind and Visually Impaired Travelers	Crabb, Ryan and Cheraghi, Seyed Ali and Coughlan, James M	Sensors	2023	Forward Snowballing
[82]	Wearable system to guide crosswalk navigation for people with visual impairment	Son, Hojun and Weiland, James	Frontiers in Electronics	2022	Forward Snowballing
[83]	Indoor Low Cost Assistive Device using 2D SLAM Based on LiDAR for Visually Impaired People	Hakim, Heba and Fadhil, Ali	Iraqi Journal for Electrical & Electronic Engineering	2019	Forward Snowballing

TABLE 10. List of conference papers included in SLR - part 1.

Ref.	Title	Authors	Published	Year	Source
[84]	Efficient Real-Time Localization in Prior Indoor Maps Using Semantic SLAM	Goswami, R. G. and Amith, P. V. and Hari, J. and Dhaygude, A. and Krishnamurthy, P. and Rizzo, J. and Tzes, A. and Khorrami, F.	9th Inter. Conf. on Automation, Robotics and Applications (ICARA)	2023	IEEE Digital Library
[85]	Detect and Approach: Close-Range Navigation Support for People with Blindness and Low Vision	Hao, Yu and Feng, Junchi and Rizzo, John-Ross and Wang, Yao and Fang, Yi	European Conf. on Computer Vision	2022	Springer Link
[87]	PathFinder: Designing a Map-Less Navigation System for Blind People in Unfamiliar Buildings	Kuribayashi, Masaki and Ishihara, Tatsuya and Sato, Daisuke and Vongkulbhisal, Jayakorn and Ram, Karnik and Kayukawa, Seita and Takagi, Hironobu and Morishima, Shigeo and Asakawa, Chieko	CHI Conf. on Human Factors in Computing Systems	2023	ACM Digital Library
[88]	A Novel Perceptive Robotic Cane with Haptic Navigation for Enabling Vision-Independent Participation in the Social Dynamics of Seat Choice	Agrawal, Shivendra and West, Mary Etta and Hayes, Bradley	IEEE Inter. Conf. on Intelligent Robots and Systems	2022	Scopus
[89]	A Multi-Sensory Blind Guidance System Based on YOLO and ORB-SLAM	Rui, Chufan and Liu, Yichen and Shen, Junru and Li, Zhaobin and Xie, Zaipeng	IEEE Inter. Conf. on Progress in Informatics and Computing (PIC)	2021	IEEE Digital Library
[90]	Indoor Navigation Assistance for Visually Impaired People via Dynamic SLAM and Panoptic Segmentation with an RGB-D Sensor	Ou, Wenyan and Zhang, Jiaming and Peng, Kunyu and Yang, Kailun and Jaworek, Gerhard and Müller, Karin and Stiefelhagen, Rainer	Inter. Conf. on Computers Helping People with Special Needs	2022	Scopus
[91]	A Wearable Robotic Device for Assistive Navigation and Object Manipulation	Jin, Lingqiu and Zhang, He and Ye, Cang	IEEE Inter. Conf. on Intelligent Robots and Systems	2021	Scopus
[92]	Multi-functional smart E-glasses for vision-based indoor navigation	Xu, Jiaqi and Xia, Haisheng and Liu, Yueyue and Li, Zhijun	Inter. Conf. on Advanced Robotics and Mechatronics (ICARM)	2021	Scopus
[93]	Personalized Navigation that Links Speaker's Ambiguous Descriptions to Indoor Objects for Low Vision People	Lu, Jun-Li and Osone, Hiroyuki and Shitara, Akihisa and Iijima, Ryo and Ryskeldiev, Bektur and Sarcar, Sayan and Ochiai, Yoichi	Inter. Conf. on Human-Computer Interaction	2021	Springer Link
[95]	Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians	Kayukawa, Seita and Ishihara, Tatsuya and Takagi, Hironobu and Morishima, Shigeo and Asakawa, Chieko	Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.	2020	ACM Digital Library
[96]	A Navigation Aid for Blind People Based on Visual Simultaneous Localization and Mapping	Chen, Cing-Han and Wang, Chien-Chun and Lin, Sian-Fong	IEEE Inter. Conf. on Consumer Electronics	2020	IEEE Digital Library
[97]	Can we unify perception and localization in assisted navigation? an indoor semantic visual positioning system for visually impaired people	Chen, Haoye and Zhang, Yingzhi and Yang, Kailun and Martinez, Manuel and Müller, Karin and Stiefelhagen, Rainer	Computers Helping People with Special Needs: 17th Inter. Conf., ICCHP	2020	Scopus
[98]	Indoor Localization for Visually Impaired Travelers Using Computer Vision on a Smartphone	Fusco, Giovanni and Coughlan, James M.	17th Inter. web for all Conf.	2020	ACM Digital Library
[99]	Human-Robot Interaction for Assisted Wayfinding of a Robotic Navigation Aid for the Blind	Zhang, He and Ye, Cang	Inter. Conf. on Human System Interaction, HSI	2019	Scopus
[100]	A Multi-Sensor Fusion System for Improving Indoor Mobility of the Visually Impaired	Zhao, Yu and Huang, Ran and Hu, Biao	Chinese Automation Congress (CAC)	2019	IEEE Digital Library
[101]	Navigation Agents for the Visually Impaired: A Sidewalk Simulator and Experiments	Weiss, Martin and Chamorro, Simon and Girgis, Roger and Luck, Margaux and Kahou, Samira E. and Cohen, Joseph P. and Nowrouzezahrai, Derek and Precup, Doina and Golemo, Florian and Pal, Chris	Proceedings of Machine Learning Research	2019	Scopus
[102]	Real-time localization and navigation in an indoor environment using monocular camera for visually impaired	Ramesh, Kruthika and Nagananda, S. N. and Ramasangu, Hariharan and Deshpande, Rohini	5th Inter. Conf. on Industrial Engineering and Applications (ICIEA)	2018	IEEE Digital Library

TABLE 11. List of conference papers included in SLR - part 2.

Ref.	Title	Authors	Published	Year	Source
[103]	Indoor Navigation using Text Extraction	Eden, Jake and Kawchak, Thomas and Narayanan, Vijaykrishnan	IEEE Inter. Workshop on Signal Processing Systems (SiPS)	2018	IEEE Digital Library
[104]	Autonomous Scooter Navigation for People with Mobility Challenges	Mulky, Rajath Swaroop and Koganti, Supradeep and Shahi, Sneha and Liu, Kaikai	IEEE Inter. Conf. on Cognitive Computing (ICCC)	2018	IEEE Digital Library
[105]	Localizing people in crosswalks using visual odometry: Preliminary results	Lalonde, Marc and St-Charles, Pierre-Luc and Loupias, Délia and Chapdelaine, Claude and Foucher, Samuel	Inter. Conf. on Pattern Recognition Applications and Methods (ICPRAM)	2018	Scopus
[106]	Plane-aided visual-inertial odometry for pose estimation of a 3D camera based indoor blind navigation	Zhang, He and Ye, Cang	British Machine Vision Conf. (BMVC)	2017	Scopus
[107]	CCNY Smart Cane	Chen, Qingtian and Khan, Muhammad and Tsangouri, Christina and Yang, Christopher and Li, Bing and Xiao, Jizhong and Zhu, Zhigang	IEEE 7th Annual Inter. Conf. on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)	2017	IEEE Digital Library
[108]	A Cloud and Vision-Based Navigation System Used for Blind People	Bai, Jinqiang and Liu, Dijun and Su, Guobin and Fu, Zhongliang	Inter. Conf. on artificial intelligence, automation and control technologies	2017	ACM Digital Library
[109]	Indoor positioning and obstacle detection for visually impaired navigation system based on LSD-SLAM	Endo, Yuki and Sato, Kei and Yamashita, Akihiro and Matsubayashi, Katsushi	Inter. Conf. on Biometrics and Kansei Engineering (ICBAKE)	2017	Scopus
[110]	SeeWay: Vision-Language Assistive Navigation for the Visually Impaired	Yang, Zongming and Yang, Liang and Kong, Liren and Wei, Ailin and Leaman, Jesse and Brooks, Johnell and Li, Bing	IEEE Inter. Conf. on Systems, Man, and Cybernetics (SMC)	2022	Forward Snowballing
[111]	The Methods of Visually Impaired Navigating and Obstacle Avoidance	Shahani, Siddharth and Gupta, Nitin	Inter. Conf. on Applied Intelligence and Sustainable Computing (ICAISC)	2023	Forward Snowballing
[112]	The Design of Person Carrier Robot using SLAM and Robust Salient Detection	Yun, Youngjae and Gwon, Taeyang and Kim, Donghan	18th Inter. Conf. on Control, Automation and Systems (ICCAS)	2018	Forward Snowballing

III. RESULT

In this section, the findings of SLR are presented. Figure 4 shows the number of studies included in this review, which focused on BVI navigation using SLAM techniques. As shown in the figure, although only papers published in the first half of 2023 are included in this review, they constitute a substantial portion of the total. The growth in the number of studies in this domain suggests an advancement in SLAM techniques and an increase in their usage for developing navigation technologies for visually impaired individuals. This section is divided into four parts to answer the research questions. It discusses the types of SLAM techniques used to develop assistive technologies for visually impaired navigation, delves into the advantages and limitations of these techniques, highlights the challenging scenarios addressed, and presents the attributes of the SLAM technology that contribute to the enhancement of visually impaired navigation.

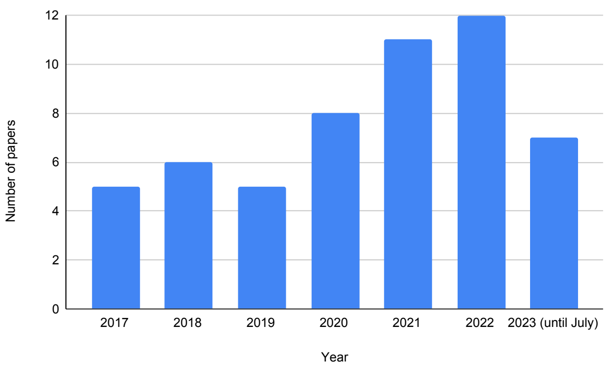

FIGURE 4. Publications included in this review on SLAM-based BVI navigation, by year

A. RQ1. WHAT LOCALIZATION AND MAPPING APPROACHES ARE USED FOR THE NAVIGATION OF THE VISUALLY IMPAIRED?

Given the variety of SLAM systems designed with different sensors, applications, and scenarios, this section focuses specifically on reviewing the types of SLAM used for the navigation of the visually impaired. It is a key technology in robotics and computer vision and has the potential to assist visually impaired individuals with navigation. This can help visually impaired individuals provide real-time location information, maps, and spatial awareness. Among the 54 studies surveyed, three strategies were common, as shown in Table 16. In this table, we use the exact terms mentioned in the literature for the localization and mapping techniques.

To further understand the technical features employed in these solutions, the detailed information is presented in Tables 17–19. These tables focus on the localization and mapping components of the assistive system, specifically highlighting the sensor types, computing resources, and application of machine learning-based methods. By examining these features, we can gain deeper insight into how these systems are structured and the diverse technologies utilized to achieve accurate and efficient SLAM for assistive navigation. It is important to note that this information relates only to the localization and mapping components of the assistive navigation solutions. Details of the entire system are provided in Tables 34–38.

This section is divided into three subsections, where we discuss the localization and mapping approaches, the sensor types used for these approaches, and the computing resources required to perform these approaches.

1) Approaches

The majority of studies have leveraged established SLAM techniques, such as ORB-SLAM, while some studies have developed new solutions tailored to their needs. For example, [73] proposed visual simultaneous localization and mapping for the moving-person tracking (VSLAMMPT) method. The proposed method was designed to assist people with disabilities, particularly visually impaired individuals, in navigating indoor environments. Additionally, various studies have used the SLAM components of existing frameworks, such as ARCore and ZED camera SLAM.

It is worth noting that several studies have employed VIO and SLAM as the core components in their proposed systems, whereas others have employed them to enhance the robustness of localization [71], for comparison with alternative localization approaches [68], or to develop new localization methods [80]. For example, [80] presented a novel localization approach specifically designed for individuals with visual impairment. This method combines visual landmark identification, VIO, and spatial constraints derived from a two-dimensional (2D) floor plan.

SLAM methodologies can be categorized into feature-based, direct, and optical-flow techniques. Feature-based methods extract and describe feature points in an image, which are then matched across different images for tracking and mapping. On the other hand, direct methods directly calculate the luminosity changes of pixel blocks. Optical flow methods utilize the optical flow changes in feature points, pixel gradient points, or the entire image to track and map the environment [66].

SLAM algorithms are further categorized into optimization- and filtering-based methods, each with distinct approaches to map creation and agent localization. Optimization-based, often referred to as Graph SLAM, treats the problem as a large optimization task, where the goal is to find the set of poses and landmark positions that best explain the observed sensor measurements. This is typically achieved by constructing a graph in which nodes represent agent poses or landmarks and edges represent constraints or observations between them. The solution is determined by minimizing the global cost function, which represents the error between the predicted and actual measurements, using nonlinear optimization techniques. On the other hand, filtering-based SLAM uses recursive Bayesian filters, such as the Extended Kalman Filter (EKF) or Particle Filter, to incrementally update the map and the agent's position as new sensor data arrive.

TABLE 12. A summary of perspectives and innovations in SLAM-based navigation solutions, with insights into limitations and advantages - Part I.

Ref.	Open-source	Research Issue	Contribution	Advantages	Limitations
[56]	•	Addressing sonification techniques for navigation instruction	Innovative sonification for BVI navigation	Publicly available localization approach	Prone to drift, utilizing markers for user positioning
[57]	•	Enhancing indoor navigation for BVI with wearable technologies	Low-cost wearable, SLAM navigation, multitarget recognition	Non-intrusive wearable device, Map reuse,	Network impact, cognitive feedback needs
[58]	•	Improving perception and independent task completion for BVI	Mixed reality, real-time perception, remote assistance	Improves perception, functions in diverse indoors, advanced interactive platform, sensor integration	Network status influence
[59]	•	Addressing guide robots' lack of consideration for user status and obstacle properties	Introduced considerate robot design and spatial risk map for navigation	Robot adapts speed and movement, enhancing natural interaction without disturbing others; calculates pedestrian directions	Needs improved speech recognition, real-world testing with BVI, effective environmental interaction
[60]	•	Addressing limitations of sensor-based navigation for BVI	Infrastructure-independent vision-based navigation for BVI	Map-evolution feedback loop ensures dynamic updates, and offline computation allows for continued use during signal loss.	Requires a dense reference image database; difficult to adapt to diverse environments.
[61]	•	Improving high-precision indoor localization in complex multi-floor environments	Hybrid localization framework combining visual and wireless signals for high-precision indoor localization	Multi-modal sensor integration, fusion-based localization, and GAN-based approach for efficient, high-precision multi-floor localization	Implementation complexity, low positioning accuracy, and dependence on fingerprint database and offline maps
[62]	•	Applying object-goal navigation to aid BVI in indoor settings	Migration of object-goal navigation to assistive devices and a knowledge-driven approach	Active assistance with context-aware, knowledge-driven navigation for improved indoor object-goal guidance	Dependency on unlabelled images, scene understanding complexity, and generalization to unfamiliar environments.
[63]	•	Operationalizing Web 2.0 social networking for indoor navigation assistance	Innovative integration of Crowdsourcing and social networking for indoor navigation	Constant access to a 24/7 indoor route database, offline functionality, and flexible, user-friendly wearable devices	Relies heavily on social networking, high processing power, and stable Internet connectivity
[64]	•	Enhancing guidance systems for BVI with multi-sensory integration for improved indoor navigation	Integration of ORB-SLAM with YOLO, dense map generation, and practical prototype implementation	Obstacle avoidance with multi-sensory feedback, dense navigation maps, and real-time target detection, implemetation of a practical smart cane	Dense map may not always align with reality, pathfinding has high computational costs, and target detection is trained based on a generic dataset.
[65]	Trajectory dataset [1]	Forecasting egocentric camera wearers' trajectories in crowded spaces	A new egocentric dataset, a Transformer-based model with cascaded cross-attention, and demonstration of socially compliant robot navigation	An egocentric view with multi-modal fusion for trajectory forecasting; Socially compliant robot navigation and assists visually impaired individuals.	Not mentioned
[66]	•	limited navigation options, wearable devices, semantic visual SLAM, cost-effective solutions	Semantic SLAM integration, real-time solution, efficient resource allocation	Real-time semantic understanding, efficient resource allocation, enhanced navigation accuracy, voice broadcast for destination assistance	Remote server challenges including internet access dependency, security, and privacy concerns
[67]	✓	Decreasing cognitive burden, increasing walking speed with a multimodal augmented cane	Improving mobility with sensors, intuitive steering, and advanced navigation features, increasing walking speed	Enhanced mobility, faster walking, precise steering, reduced cognitive load, fewer collisions, increased confidence, reliable obstacle avoidance, failure backup, selecting preferred walking speed by user	Heavy, requiring mechanical assembly
[68]	•	Enhancing navigation with improved robotic assistance in dynamic environments.	A novel haptic-guided robot with enhanced navigation, dynamic obstacle avoidance via UWB, and voice-enabled beacon feedback	Enhanced environmental information, intuitive navigation, precise UWB positioning, semantic feedback, and DRL-based obstacle avoidance.	Dynamic obstacles impact SLAM accuracy; implicit learning in the simulator may cause real-world uncertainties.
[69]	•	Efficient indoor navigation aid.	Multi sensor utilization, efficient algorithm utilization, and real-time voice guidance.	Cost-effective, real-time navigation, accurate mapping, obstacle detection, efficient path planning.	Limited to static, simple environments.
[70]	•	Enhancing navigation accuracy and robustness in large indoor spaces for assistive technologies.	Introduced DVIO for enhanced 6-DOF pose estimation and VPS for accurate assistive navigation.	Enhanced pose accuracy, real-time updates, effective wayfinding, obstacle avoidance, and significant pose error reduction.	Pose drift remains an issue in complex environments.

[1] https://github.com/Jianing-Qiu/TISS (Accessed on 3 May 2024)

Systems such as ORB-SLAM and RTAB-Map are feature- and optimization-based, employing features and graph optimization for mapping and localization. Conversely, LSD-SLAM and DSO are examples of direct and optimization-based SLAM. Some systems, such as Semantic SLAM, may adopt either approach, depending on their implementation. It is important to note that the method and type of SLAMs are not directly mentioned in all papers, so the information provided here is a general categorization based on common practices within each category.

The reviewed studies demonstrate the versatility of SLAM in various navigation scenarios. The specific implementation and aspects addressed in each study varied depending on the application. SLAM can be employed in environments that lack a map and can dynamically create it while navigating. This involves simultaneous map creation and localization within an environment. Alternatively, SLAM can be employed to generate maps for subsequent navigation. In this case, SLAM first builds a map of the environment and then the map is utilized during navigation. Studies have also used SLAM odometry for navigational tracking. Odometry provides a continuous estimate of the position and orientation

TABLE 13. A summary of perspectives and innovations in SLAM-based navigation solutions, with insights into limitations and advantages - Part II.

Ref.	Open-source	Research Issue	Contribution	Advantages	Limitations
[71]	•	Improving visual localization in challenging outdoor settings.	Unified Dual Desc network enhancing descriptor extraction and multi-modal integration for assistive localization.	Enhances robustness using multi-modal imaging, advanced descriptors, and sequential integration for outdoor navigation.	Lacks real-time execution capability in assistive device.
[72]	•	Enhancing independence for visually impaired through semantic environmental understanding and navigation.	Integrates ConvNets for semantic mapping; enhances topological localization.	Enhanced accuracy, semantic guidance, robustness to changes.	Computational inefficiency, suboptimal performance in motion blur effects.
[73]	•	Addressing the lack of comprehensive solutions for indoor navigation and obstacle avoidance.	Integration of dynamic person-detection method (EER–ASSL) and VSLAM for real-time navigation assistance in cluttered environments.	Enhanced smooth movement, reliable obstacle avoidance, effective navigation in dynamic environments.	Limited instruction capabilities, the decrease in person detection performance under varying lighting conditions and speed.
[74]	•	Addressing lack of integrated navigation and object recognition systems.	Integration of lightweight CNN-based object recognition and visual SLAM for improved environment perception and navigation.	Load cognitive decrease, safe and quick navigation, enhanced perception, and real-time performance on smartphones.	Unable to detect small-size obstacles and obstacle detection limitations.
[75]	•	Addressing lack of user-centric indoor navigation aids for visually impaired.	ARCore integration, adaptive path planning, and dual-channel user interaction for indoor navigation.	Enhanced mapping, obstacle-avoiding path planning, and intuitive dual-channel interaction for improved indoor navigation.	Reliance on existing indoor scenario CAD maps.
[76]	•	Addressing gaps in localization, way-finding, and route following for visually impaired navigation.	Dynamic subgoal route following, visual SLAM integration, and wearable optical see-through glasses for enhanced indoor navigation.	Enhances precision with visual SLAM, cost-effective sensors, efficient obstacle avoidance, safe indoor navigation with dynamic subgoal selection.	Not mentioned
[77]	•	Addressing accumulative pose error, GPS-denied navigation, and real-time pose estimation.	New 6-DOF pose estimation method using floor and wall information, and a real-time wayfinding system.	Reduces accumulative pose error and provides real-time wayfinding	System less effective in simple tasks, weight causes discomfort, and fails at high walking speeds over 0.6 m/s.
[78]	•	Addressing accurate 6-DOF pose estimation challenges through innovative visual-inertial odometry for robotic navigation aids.	A plane-aided VIO method and a plane-consistency check for enhanced pose estimation accuracy.	Improved accuracy, a plane-consistency check, practical implementation for assistive navigation, and outperformance of state-of-the-art methods.	Not mentioned
[79]	•	Enhancing navigation in an indoor public environment for pre-scheduled tasks with user-independent robotic solutions.	Innovative user-independent robotic assistance for indoor navigation.	Enhancing inclusivity and efficiency: user-independent robotic assistance.	It is not mentioned how the proposed approach handles crowds in the public places under study.
[80]	ICOSR repository [2]	Developing a lightweight indoor localization system using a 2D floor plan of the environment rather than a 3D model.	Innovative localization algorithm integrating visual landmarks, VIO, and 2D floor plans.	Smartphone-based, lightweight, and robust localization approach.	Noisy distance estimates due to imprecise bounding boxes.
[82]	Raw data supporting the conclusion	Addressing the need to ensure safe street crossing	Introducing a comprehensive wearable system for safe urban navigation, integrating real-time computer vision and prior maps.	Utilizes pre-built LiDAR maps, such as those publicly available and created for autonomous vehicles, supports user's preferred walking speed.	Dependent on specific crosswalk textures, lacks dynamic obstacle handling, may face instability with changing features, and outdoor noise interference.
[83]	•	Integration of navigation and object recognition.	Integration of navigation, object recognition, and low-cost sensors.	Integrated navigation and object recognition, low-cost sensors, efficient path planning, accurate real-time object identification in static scenarios.	Restricted to static simple environments.
[84]	•	Enhancing real-time global indoor localization using Semantic SLAM and a priori maps for GPS-deprived environments.	Implementing vector-based semantic extraction from floor plans, efficient particle filter localization, and leveraging loop closures for active semantic point cloud.	Real-time, accurate localization; integration of semantic information; using deep learning for enhanced robustness and efficiency.	Visual aliasing, limited semantic classes.
[85]	•	Addressing navigation challenges for BVI by developing a wearable solution for real-time guidance to target objects in unfamiliar settings.	Introducing a wearable navigation system with real-time object localization, visual SLAM, and trajectory estimation for efficient user guidance.	High accuracy, vision-based, real-time assistance, continuous object tracking, and portable system design.	Not explicitly mentioned.

[2] https://www.openicpsr.org/openicpsr/project/183714/version/V1/view (Accessed on 3 May 2024)

of an agent based on the sensor readings.

Our analysis, underscored by the classifications in Table 16, indicates a strong preference for feature-based and optimization-based SLAM approaches for visually impaired navigation. This preference is likely due to the robustness and efficiency of these methods in processing visual data, which is key for real-time assistive navigation.

Figure 5 provides insight into the use of various localization and mapping techniques for visually impaired navigation from 2017 to the date when SLR was conducted (July 2023). This figure illustrates that visual techniques has consistently been used across all years. Although many other techniques also operate based on visual data, we mentioned each of these techniques, as indicated in the referenced studies.

The utilization of semantic SLAM and Cartographer SLAM signifies a recent trend towards leveraging advanced spatial understanding and mapping capabilities for visually impaired navigation. Semantic SLAM incorporates higher-level scene interpretation and enhances users' contextual awareness. On the other hand, Cartographer SLAM provides

TABLE 14. A summary of perspectives and innovations in SLAM-based navigation solutions, with insights into limitations and advantages - Part III.

Ref.	Open-source	Research Issue	Contribution	Advantages	Limitations
[87]	•	Addressing the challenge of navigation in unfamiliar indoor spaces by designing a map-less system, PathFinder	Developing a map-less navigation robot system incorporating sign recognition and intersection detection, using scenario-based participatory design with five blind participants.	Enhanced navigation confidence, provided key sign information, offered audio feedback, ensured safe, independent mobility by PathFinder	Limited study environments, assumed environments without steps or floor transitions, no empirical comparison with participants' regular aids, inability to recruit younger participants and insufficient number of guide dog users, possible positive bias from participants who have previously participated in studies, physical demand, Bluetooth connectivity issues, privacy concerns, inability to navigate in congested spaces, surrounding people may not recognize that users are disabled as they don't use traditional aids.
[88]	•	Enhancing independent navigation and social seat selection with a perceptive robotic cane system.	Introducing a robotic cane with computer vision for navigation and social seat selection, featuring vibrotactile feedback and successful pilot validation.	Independent navigation, social norm-aware seat selection, intuitive vibrotactile feedback, and effective pilot-validated guidance.	Potential suboptimal chair detection, discrepancy between user preferences and performance.
[89]	•	Overcoming limitations of existing blind navigation methods by integrating multi-sensory feedback for comprehensive and intuitive mobility.	Integration of YOLO and ORB-SLAM, enhanced by novel algorithms, provides reliable multi-sensory guidance	Enhanced accuracy, multi-sensory feedback, real-time object detection, dense navigation maps, and comprehensive guidance improve mobility and safety.	Not explicitly mentioned
[90]	•	Addressing gap: aiding visually impaired in dynamic indoor navigation, obstacle detection, and social distancing.	Wearable RGB-D assistant system aiding indoor localization, mapping, and dynamic obstacle detection.	Dynamic object detection, enhanced obstacle avoidance, panoptic segmentation for scene understanding, robust tracking without additional models, and RGB-D sensor integration.	Imperfect panoptic segmentation, leading to errors in object recognition; computational complexity, influenced by the number of dynamic objects present in the scene.
[91]	•	Addressing the gap: wearable device aiding visually impaired with indoor object manipulation tasks.	Hand-worn assistive device, RGBD-VIO method, effective human-device interface.	Enhanced pose estimation, depth information utilization, improved human-device interaction, effective object manipulation.	Not mentioned
[92]	•	Addressing the challenge of indoor navigation, integrating SLAM and deep learning enhances environmental perception.	Introducing multi-functional smart E-Glasses for enhanced indoor navigation and lightweight object detection.	Real-time navigation, high object detection precision, and robust SLAM integration.	Future enhancements aim to improve device comfort, portability, and size, addressing user concerns and enhancing overall usability.
[93]	•	Addressing the gap in understanding between visually impaired individuals' perceptions and their actual surroundings for navigation technology.	Introducing personalized navigation system with object detection and description, and re-training models.	Personalized navigation, object detection, reduced time for finding destinations, and improved interaction.	Model re-training challenges, environmental complexity, and high computational cost.
[95]	•	Addressing the need for a guiding system to facilitate seamless walking in public spaces.	Introducing a comprehensive system aiding blind pedestrians by understanding nearby pedestrians' behavior.	Convenient suitcase design, accurate motion tracking, effective tactile interface for enhanced navigation.	Weight discomfort, space constraints, speed adjustment difficulties, and technological improvements for enhanced usability.
[96]	•	Addressing the need for advanced navigation aid for the visually impaired using VSLAM technology.	Introducing a navigation aid system merging VSLAM with pre-established maps.	Not explicitly mentioned	Not explicitly mentioned
[97]	•	Addressing the need for unified indoor navigation, integrating scene perception and visual localization.	Introducing a unified semantic visual localization system, enhancing obstacle avoidance and spatial awareness.	Real-time awareness, comprehensive understanding, and obstacle avoidance.	Restricted camera field of view and inconsistencies in semantic segmentation results impacting user confidence.
[98]	Upon publication	Addressing indoor navigation challenges through smartphone-based computer vision without new infrastructure.	Real-time app development, robust localization algorithm, and user-friendly navigation.	Cost-effective deployment, enhances usability, improves localization accuracy, promises full-featured wayfinding, and camera-agnostic navigation for ease of use.	Challenges in wide, open indoor spaces due to limitations in current approach, suggesting potential for augmented reality integration.

SLAM in 2D and 3D across various platforms and sensor configurations, offering innovative solutions to tackle the diverse challenges associated with BVI navigation.

ORB-SLAM algorithms, including ORB-SLAM (published in 2015), ORB-SLAM2 (published in 2017), and ORB-SLAM3 (published in 2021), have gained popularity because of their robustness and performance. This can be attributed to its efficient feature extraction and matching techniques, making it well suited for real-time navigation applications.

Customized techniques have been developed to meet specific needs. This trend indicates that researchers have adjusted the SLAM techniques to better match the specific requirements of their intended applications. This suggests closer integration of SLAM with domain-specific needs.

2) Sensor type

The sensors employed in SLAM solutions for BVI navigation are diverse and include various types of cameras, LiDAR, IMU, and other specialized sensors. As shown in Figure 6, we categorized the sensors into three types: cameras, LiDAR, and other sensors.

TABLE 15. A summary of perspectives and innovations in SLAM-based navigation solutions, with insights into limitations and advantages - Part IV.

Ref.	Open-source	Research Issue	Contribution	Advantages	Limitations
[99]	•	Addressing the indoor wayfinding problem	Introducing specialized robotic navigation aid, VIO method, guiding modes, and Human Intent Detection for enhanced navigation assistance.	Enhanced VIO method, two guiding modes, and automated mode selection.	Validation in larger spaces, user feedback incorporation, and operational restrictions.
[100]	•	Addressing independent indoor corridor navigation through multi-sensor fusion and semantic mapping.	Semantic mapping, multi-sensor fusion, real-time performance enhancement	Enhanced corridor navigation, semantic mapping, landmark detection, real-time performance, multi-sensor fusion for improved navigation experience.	Object recognition limitations, arrow direction detection issue, training requirement for unknown objects.
[101]	SEVN-data [3]	Enhancing Reinforcement Learning environments to develop a navigation assistant tailored for the BVI community.	Developing a benchmark dataset and Reinforcement Learning training environment to advance navigation agent capabilities using real-world imagery and neural architecture.	SEVN offers realistic training with a rich, annotated dataset and a multi-modal fusion model for effective BVI navigation.	SEVN offers an extensible Reinforcement Learning environment, but improved model performance requires.
[102]	•	Addressing real-time localization and navigation in indoor settings with a monocular camera, focusing on computational efficiency, user-friendly interfaces, and integrated algorithms.	Introducing a non-filter based visual SLAM with integrated object detection and distance-depth estimation algorithms, using a single monocular camera for BVI indoor navigation.	Utilizing a single camera and simpler SLAM algorithm for cost-effective, efficient real-time performance.	Not mentioned
[103]	•	Enhancing indoor localization for visually impaired shoppers using text extraction, addressing gaps in traditional visual assistance systems.	Expanding SLAM algorithm for larger spaces using GIST/SURF features and navigating through text-rich environments.	Simplistic setup, no markers needed, and efficient real-time localization.	Operates only within length dimension of the aisle, limited by text density.
[104]	Android APIs	Enabling safe autonomous navigation for elderly and visually impaired in crowded environments.	Design of an intelligent autonomous scooter with advanced sensor fusion, SLAM techniques, and hybrid mapping solutions.	Improved safety and autonomy, hybrid mapping for diverse environments, precise steering control.	Not mentioned
[105]	•	Localizing pedestrians in crosswalks using visual odometry, addressing challenges in uniform textures and repetitive landmarks.	Introducing a prototype for localizing pedestrians in crosswalks using visual odometry.	Accurate on weakly textured surfaces, addressing scaling issues in monocular camera setups.	Initialization issues, tracking loss due to strong orientation variations, and challenges with oscillatory walking patterns.
[106]	•	Improving indoor pose estimation accuracy for navigation aids using plane features in feature-sparse environments.	Introducing PAVIO method, utilizing plane features and factor graph optimization to improve pose estimation for indoor navigation.	Improved pose estimation accuracy and robustness, enhanced stability, accurate 3D mapping.	Not mentioned
[107]	•	Addressing indoor navigation challenges to enhance mobility and independence.	Implementation of SmartCane with Google Tango for real-time indoor navigation and demonstration of its effectiveness.	Enhanced indoor navigation with real-time path planning, multimodal feedback, and an intuitive control panel interface.	Requiring further user evaluation to assess effectiveness
[108]	•	Developing a cloud and vision-based system for safe navigation and detailed perception.	Integrating cloud computing with vision-based navigation, enhancing perception, and improving object recognition for blind individuals.	Detailed perception, real-person safety support, abundant surrounding information, and improved object recognition.	Requiring extensive vision-based mapping, struggles with similar scenarios, and needs improved scene parsing, currency validation, and object recognition.
[109]	•	Addressing the need for specialized navigation systems for visually impaired individuals using SLAM technology for real-time guidance.	Introducing a wearable camera with LSD-SLAM for real-time positioning, obstacle detection, and route guidance.	Efficient calculation power, robust performance, accurate mapping, dynamic adaptation, real-time assistance.	Requires high-contrast environments for accurate mapping; low contrast may need external positioning solutions.
[110]	•	Addressing the need for an innovative navigation system using vision-language model-based approach.	First BVI navigation system using spoken instructions, visual-language integration, and heuristic-based path planning for improved success.	Running on portable devices, provides BVI navigation without heavy labeling or 3D model reconstruction in complex indoor environments.	Navigation reliability drops for long distances.
[111]	•	Addressing real-time navigation latency, safe route selection, and accurate obstacle detection.	Integration of Web of Things, predictive analytics, YOLOv4 Tiny, and SLAM for enhanced obstacle recognition and navigation.	Enhanced obstacle recognition, navigation, and safe route selection.	Requiring high-contrast environments; needing external updates when contrast is insufficient for accurate mapping.
[112]	•	Addressing the need for efficient navigation solutions for visually impaired individuals and patients with lower body injuries indoors.	Developed a person carrier robot integrating Hector SLAM and Robust Salient Detection for safe navigation and obstacle avoidance.	Enhanced safety, improved mobility, and effective object detection for indoor navigation.	Sensitive to strong light, issues with reflective surfaces, and occasional collisions due to slow processing speed.

[3] https://github.com/mweiss17/SEVN-data (Accessed on 12 May 2024)

Camera The common use of visual sensors in SLAM techniques can be attributed to advances in computer vision and image processing, which enhance navigation capabilities by providing rich environmental information. This makes visual-sensor-based SLAM techniques the most commonly used in the implementation of assistive technologies for the BVI people, offering a cost-effective, versatile, and accurate solution for real-time navigation and spatial awareness. The literature under review used the following types of cameras: RGB, RGB-D, stereo, monocular, and other specialized cameras.

RGB cameras are widely used due to their ability to capture rich color information, which is beneficial for visual odometry and object recognition. They are cost effective and widely available, making them a popular choice for the development of accessible navigation aids.

RGB-D cameras provide both color and depth information, enabling more accurate mapping and localization. Depth information helps in understanding the 3D structure of the environment.

TABLE 16. Localization and mapping approaches used for visually impaired navigation, with "NA" denoting data not available.

Technique	Type	Method	Sensor type	Reference(s)
Widely applicable techniques				
Cartographer	Scan matching	Optimization-based	LiDAR	[59], [61], [87], [95]
Hector SLAM	Scan matching	Optimization-based	LiDAR	[69], [83], [112]
LiDAR SLAM	Scan matching	Rao-Blackwellized particle filter	LiDAR	[67], [68], [79]
FastSLAM [121]	Feature-based	Particle filter-based	LiDAR	[82]: mapping
Kinect Fusion	Feature-based	Optimization-based	Visual	[58]
LSD-SLAM	Direct	Optimization-based	Visual	[109], [111]
OpenVSLAM	Feature-based	Optimization-based	Visual	[60], [93], [97]
ORB-SLAM	Feature-based	Optimization-based	Visual	[64], [89]
ORB-SLAM2	Feature-based	Optimization-based	Visual	[57], [62], [76], [92], [96], [101], [82]: localization
ORB-SLAM3	Feature-based	Optimization-based	Visual	[63], [65]
Pose-graph SLAM	Feature-based	Optimization-based	Visual	[77]
RTAB-Map	Feature-based	Optimization-based	Visual	[104]
Semantic SLAM	Feature-based [66], [84] , NA [100]	Optimization-based [66], Particle filter-based [84], NA [100]	Visual	[66], [84], [100]
Visual SLAM	Feature-based	Optimization-based	Visual	[71], [72], [74], [85], [102], [108]
DSO	Direct	Optimization-based	Visual	[105]
VIO (Visual Inertial Odometry)	Feature-based	Optimization-based	Visual	[56], [98], [99]
Customized solution				
DVIO	Feature-based	Optimization-based	Visual	[70]
PAVIO	Feature-based	Optimization-based	Visual	[78], [106]
Dynamic-SLAM	Feature-based	Optimization-based	Visual	[90]
RGBD-VIO	Feature-based	Optimization-based	Visual	[91]
VSLAMMPT	Feature-based	Optimization-based	Visual	[73]
Spatial tracking frameworks				
Google ARCore	Feature-based	Likely optimization-based	Visual	[75]
ZED camera's SLAM	Likely feature-based	Not explicitly stated	Visual	[103]
Apple iOS ARKit-based	Likely feature-based	Update process is implemented using a particle filter [80]	Visual	[80], [110]
Intel RealSense SLAM	Likely feature-based	Not explicitly stated	Visual	[88]
Google Tango's built-in SLAM	Likely feature-based	Not explicitly stated	Visual	[107]

Stereo cameras also provide depth perception through two slightly offset lenses that simulate binocular vision. They are effective in capturing detailed depth information and are useful in applications where precise depth estimation is required.

Monocular cameras are simpler than stereo and RGB-D cameras. They rely on visual odometry and other techniques to estimate depth and motion, making them lightweight and suitable for mobile applications.

Specialized cameras, including fisheye, 3D time-of-flight, and wide-angle cameras, provide specialized capabilities, such as a wide field of view or precise depth measurement, which can enhance the SLAM performance in specific scenarios.

LiDAR LiDAR sensors are highly accurate in measuring distances and are effective in creating detailed 3D maps of the environment. Studies use LiDAR alone to build a map of the environment or in combination with other sensors such as IMU and cameras to enhance the robustness and accuracy of SLAM systems.

Other sensors Various studies combined different types of sensors to leverage the strengths of each type and provide more robust and reliable navigation solutions. For example, integrating an IMU with a camera helps achieve better motion tracking and stability. This trend towards integrating multiple sensors highlights increasing efforts to enhance the robustness and reliability of SLAM solutions.

Table 20 provides a detailed breakdown of the sensor types employed in the reviewed studies, highlighting the prevalence of different sensor modalities and their combinations in SLAM-based assistive technologies for BVI individuals.

The reviewed papers show a clear preference for RGB-D cameras, indicating their effectiveness in providing both the visual and depth information necessary for accurate SLAM applications. The use of LiDAR is important in applications that require precise mapping. Over the years, there has been a noticeable trend towards integrating multiple sensors and combining their strengths to achieve more robust and reliable SLAM solutions for visually impaired navigation. The integration of machine learning-based techniques with SLAM systems is particularly prevalent in solutions that utilize RGB-D cameras. This highlights the effectiveness of combining this type of data with advanced AI algorithms. This trend is likely to continue as technology advances, offering more sophisticated and adaptable solutions in complex and dynamic environments.

3) Computing resource

To process data and run localization and mapping algorithms, the reviewed studies adopted two classes of computational resources: local and remote. Local computations are performed in situ on devices, such as smartphones, tablets, laptops, portable microcontrollers, and UP board computers. In some cases, algorithms were applied on PCs. Table 21 categorizes the computing resources used in the reviewed studies for localization and mapping tasks. Information regarding the computing resources for the entire navigation assistive system

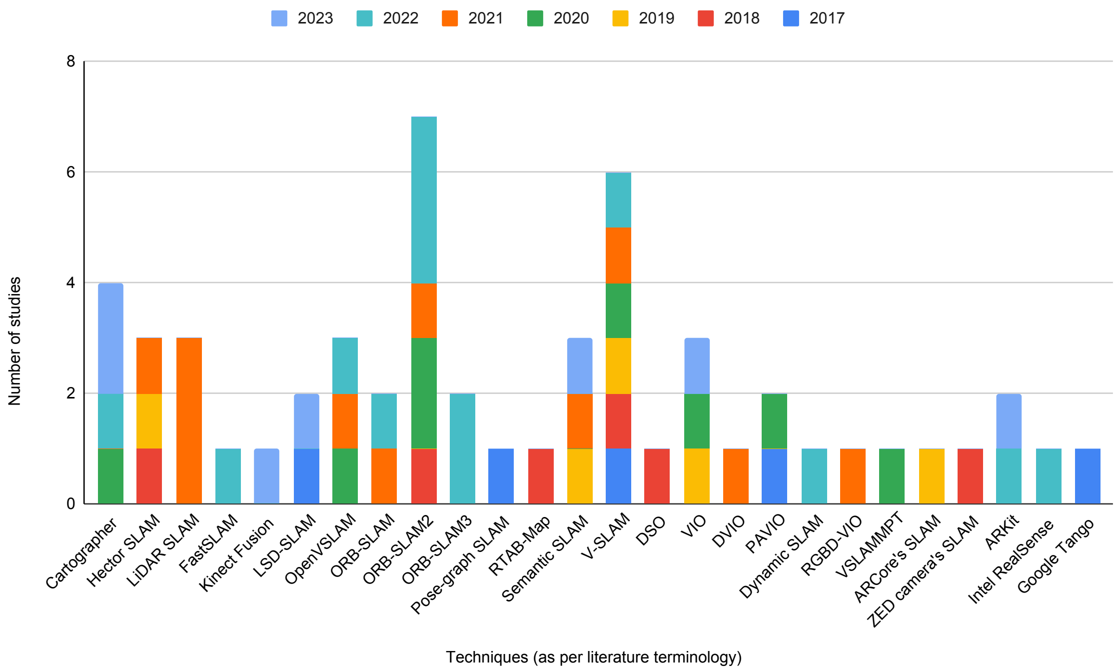

FIGURE 5. **Evolution and adoption of localization and mapping techniques in BVI navigation systems over time.**

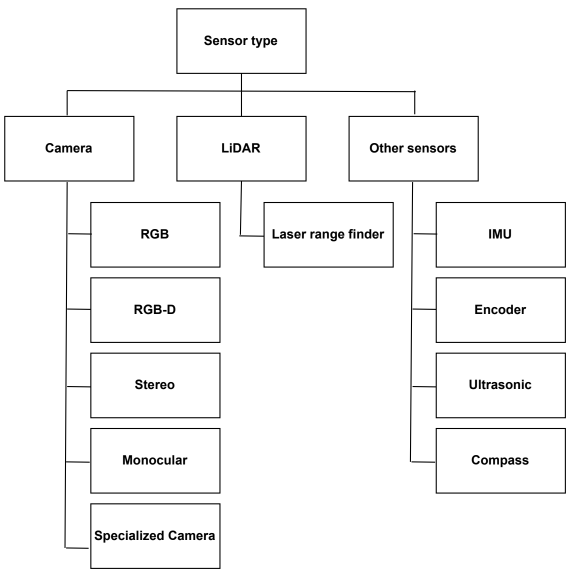

FIGURE 6. **Overview of sensor types in studies under review**

is detailed in Tables 36-38

a: Local computing resources

Smartphones: Smartphone are widely used as communication gadgets, and their technology continues to grow to the point that it is possible for smartphones to implement functional navigation systems. Because smartphones integrate diverse sensors such as IMU, GPS, and cameras, they can be used as a convenient tool for collecting environmental information. In addition, their computational power can be exploited to perform various navigation operations. For example, the system proposed by [74] implemented all algorithms relevant to data acquisition, ground segmentation, moving direction search, global path planning, indoor and outdoor localization, and object detection on a smartphone and achieved real-time performance. Without an additional depth sensor, [75] took advantage of an ARCore-supported smartphone to track pose and to build a map of the surroundings in real time. However, despite the significant advantages of smartphones, such as their small size, low weight, easy portability, and low cost, their computing power is not sufficient for some approaches.

Laptops and PCs: Some of the reviewed approaches perform all or part of the required calculations locally on a portable computer, such as a laptop. Despite higher computing power compared to a smartphone, and greater security compared to remote computational resources, the laptop's heavier weight and large size are considered major disadvantages, especially during long trips. PCs provide even higher computing power, which is essential for complex SLAM operations; however, they lack portability.

TABLE 17. Comparison of core technical features for localization and mapping techniques, with a specific focus on sensor types, computing resources, and whether machine learning-based methods are employed for localization and mapping tasks - Part I.

Ref.	Sensor type	Computing Resource	Localization & Mapping Technique	ML-Based Localization and Mapping
[56]	Camera	Smartphone	Native AI library for iOS devices	Not explicitly mentioned
[57]	RGB-D camera	Remote server	ORB-SLAM2	•
[58]	A depth, an RGB, and four gray scale cameras, an IMU	Hololens2 device, GPU	Iterative Closest Point (ICP) for camera pose estimation, Kinect Fusion algorithm for real-time 3D reconstruction	Not explicitly mentioned
[59]	Wheel encoder and LiDAR	Notebook PC	Cartographer	•
[60]	RGB camera	Cloud server and Nvidia Jetson AGX Xavier	Visual place recognition, weighted averaging, and perspective-n-point (PnP) for localization, OpenVSLAM and Colmap to generate a topometric map	NetVLAD for global descriptors and SuperPoint for local descriptors
[61]	LiDAR	Not mentioned	Cartographer to build SLAM maps on each floor	•
[62]	RGB-D camera	Jetson AGX Xavier	ORB-SLAM2 for localization; downsampling, octomap, and 2D occupancy grid mapping for an accurate dense mapping	•
[63]	IMU, stereo and IR (depth) cameras	Cloud server	ORB-SLAM3	•
[64]	RGB-D camera	Raspberry Pi	ORB-SLAM	•
[65]	Monocular RGB	Not mentioned	ORB-SLAM3 to obtain ground-truth camera trajectory	•
[66]	RGB-D camera	High performance portable processor, cloud server	Semantic visual SLAM based on ORB feature to generate sparse, dense, and semantic maps	ENet for pixel-level semantic segmentation
[67]	LiDAR	Raspberry Pi	BreezySLAM	•
[68]	LiDAR	Intel NUC computer	GMapping for mapping and estimating the destination's location	•
[69]	RGB-D camera	Raspberry Pi3 B+	Hector SLAM for building the environmental map and locating the user on the map	•
[70]	RGB-D Camera, IMU	UP Board computer	A VIO system developed based on VINS-Mono	•
[71]	RGB-D-IR camera	A portable computer, Nvidia Jetson TX2	Hierarchical visual localization pipeline with deep descriptors, geometric verification, and sequence matching.	NetVLAD and Dense Desc for advanced descriptor extraction.
[72]	RGB-D camera	Odroid XU3 board, remote server	Off-the-shelf algorithm [120] for 3D indoor mapping; a two-stream ConvNet for topological localization.	ConvNet replaces BoW for semantic info; Inception-v3 enhances object recognition, aiding localization.
[73]	Two monocular cameras	Not mentioned	Followed a similar structure as ORB-SLAM2.	•
[74]	RGB-D camera, IMU	A smartphone with Qualcomm Snapdragon 820 CPU 2.0 GHz	Vins-mono for indoor localization; ORB-SLAM2 and Vins-mono for building a indoor map repository.	•
[75]	Smartphone's camera	A HUAWEI P20 smartphone with Kirin 970 CPU	Roberto Lopez Mendez ARCore SLAM as the base for visual odometry and area learning	Not explicitly mentioned
[76]	Fisheye and depth camera	An embedded CPU board	ORB-SLAM2 for the building of the virtual-blind-road (offline by a sighted person) and the localization (online)	•
[77]	SwissRanger SR4000 3D camera	A Lenovo ThinkPad T430 laptop	2-step graph SLAM	•
[78]	SwissRanger SR4000 camera, IMU	Up Board computer	PAVIO: Fusing visual, inertial, and plane features for robust SLAM localization and mapping.	•
[79]	Encoder, IMU, laser distance sensor	Raspberry Pi 3 Model B and B+	Gmapping for building a 2D occupancy grid map on the environment	•
[80]	iPhone 11 Pro camera and IMU	iPhone 11 Pro	ARKit VIO for relative movements; ARKit mapping for semantic labels; Monte Carlo localization	YOLOv2 for object detection to facilitate effective localization

Embedded systems and microcontrollers: Embedded systems such as Nvidia Jetson boards, Raspberry Pi, and UP boards provide a balance between computational power and portability. They are commonly used in the reviewed studies for performing SLAM operations locally. For instance, [58] utilized a Hololens2 device with a GPU for real-time 3D reconstruction, whereas [62] used a Jetson AGX Xavier for ORB-SLAM2 and dense mapping. Raspberry Pi devices are also popular due to their low cost and sufficient computing power for many SLAM tasks [67], [69].

b: *Remote computing resources*

An alternative solution is to transfer all or part of the calculations to the remote computing resources. To reduce local computing costs, [57] adopted an embedded computer and a remote server. In the proposed vision-based assistance system, before transferring the input images to the server, the images were time-stamped and encrypted on an embedded computer. The remote server was equipped with a CPU and GPU to run parallel ORB-SLAM2 and artificial intelligence algorithms for indoor navigation, object detection, face recognition, and scene text recognition. Experiments confirmed that the use of remote servers under a smooth network connection, such

TABLE 18. Comparison of core technical features for localization and mapping techniques, with a specific focus on sensor types, computing resources, and whether machine learning-based methods are employed for localization and mapping tasks - Part II.

Ref.	Sensor type	Computing Resource	Localization & Mapping Technique	ML-Based Localization and Mapping
[82]	RGB-D camera, compass sensor	Jetson Xavier AGX, NVidia	FastSLAM for mapping, ORB-SLAM2 for local pose estimation in the pre-built map	An extra thread aligns predicted semantics with key features.
[83]	LiDAR, ultrasonic sensor, Raspberry pi camera	Raspberry Pi3 B+	Hector SLAM for constructing a 2D-map of the environments and localization	•
[84]	RGB-D camera, IMU	Nvidia Jetson AGX Xavier microprocessor	RTAB-Map for semantic point cloud generation and global localization	MobileNetV2 with PPM for constructing semantic point cloud
[85]	Monocular camera	Nvidia Jetson Xavier NX Developer kit	Visual SLAM for user movement estimation and stationary object localization	•
[87]	LiDAR, IMU	Nvidia RTX 3080 graphic board	Cartographer for constructing a LiDAR map	•
[88]	RGB-D camera, IMU	Dell G15 laptop with an RTX 3060 GPU	SLAM implementation by RealSense for creating an initial 2D occupancy grid and estimating user pose	•
[89]	RBG-D camera	Uzel US-M5422 edge server, Raspberry Pi 4B	Improved ORB-SLAM for generating a dense navigation map and real-time positioning	•
[90]	RGB-D camera	Laptop	Dynamic-SLAM based on ORB-SLAM2 for estimating user's ego-pose and building a static feature point map	Non-Prior dynamic object detection preceding local mapping
[91]	RGB-D camera, IMU	Google Pixel 3 smartphone	RGBD-VIO for mapping and accurately estimating the device's pose	•
[92]	RealSense D435i camera	Remote server based on Intel i7-8700 CPU, nvidia GTX1080 GPU	ORB-SLAM2 for environmental mapping and positioning of the user	•
[93]	Google glasses camera	GPU server	OpenVSLAM for mapping and locating user's indoor positions	•
[95]	LiDAR, IMU	Laptop	Cartographer for estimating the current location and direction of a user	•
[96]	RGB-D camera, IMU	Not mentioned	ORB-SLAM2 for mapping and localizing the user	•
[97]	RGB-D camera	Nvidia Jetson AGX Xavier processor	OpenVSLAM for robust mapping and localization in real time.	•
[98]	iPhone 8's IMU and rear-facing camera	iPhone 8	Visual-Inertial Odometry with sign recognition and geometric constraints.	Not explicitly mentioned
[99]	RGB-D camera, IMU	UP Board computer	Visual-Inertial Odometry for 3D mapping and pose estimation	•
[100]	RGB-D, LiDAR	Laptop	Semantic SLAM based on a 2D SLAM technique for determining the corridor area and mapping to a semantic map	YOLOv3 for landmark detection, Places365 for place recognition
[101]	Vuze+ camera	Not mentioned	ORB-SLAM2 for generating 3D positions and 2D connectivity graph from Vuze+ footage	•
[102]	Monocular camera	Intel i7 processor	Non-filter based visual SLAM using trained objects as landmarks for localization	ACF detector for object detection to identify trained objects of interest for localization.
[103]	ZED camera	Nvidia Jetson TX2	ZED camera's SLAM for navigating toward an aisle in a grocery store.	Not mentioned
[104]	Stereo camera, laser	Nvidia Jetson TX2, Raspberry Pi, and Arduino.	RTAB-MAP for obtaining fine-grained mapping of the 3D spatial world.	•
[105]	Monocular camera	Not mentioned	Visual odometry for user localization during street crossings	•
[106]	3D time-of-flight camera , IMU	UP Board computer	Plane-aided visual-inertial odometry (PAVIO) for pose estimation of an robotic navigation aid	•
[107]	Wide-angle lens camera, gyroscope, accelerometers, and infrared sensor on Google Tango.	Google Tango	Google Tango's built-in SLAM for mapping the environment and localizing the user.	•
[108]	Stereo camera	Cloud server	Visual SLAM for mapping the environment and localizing the user.	•

as 4G or WiFi, can meet the computational requirements of the proposed system. However, although the high computing power of remote servers is considered a significant advantage, constant Internet access over a secure connection is required. Moreover, the performance of the entire system would be affected by the network condition.

B. RQ2. WHAT ARE THE ADVANTAGES AND LIMITATIONS OF SLAM TECHNIQUES FOR BVI NAVIGATION?

Although SLAM techniques offer significant benefits for navigation across various applications, their use in improving mobility for BVI individuals presents unique considerations. In general applications, the advantages of SLAM include accurate real-time mapping and localization, adaptability to unknown environments, and the ability to function without an external infrastructure. These advantages are essential when applied to BVI navigation. For instance, the ability to provide real-time accurate spatial information is crucial for safe navigation and obstacle avoidance.

Conversely, some limitations of SLAM in general applications, such as computational complexity and sensor dependencies, pose significant challenges in the BVI context. The

TABLE 19. Comparison of core technical features for localization and mapping techniques, with a specific focus on sensor types, computing resources, and whether machine learning-based methods are employed for localization and mapping tasks - Part III.

Ref.	Sensor type	Computing Resource	Localization & Mapping Technique	ML-Based Localization and Mapping
[109]	Monocular camera	A single CPU	LSD-SLAM for estimating the user's position and constructing a 3D environmental map	•
[110]	Camera, IMU, LiDAR	Smartphone	visual-inertial SLAM based pose estimation and 2D scene-graph map construction using iOS ARKit.	Not explicitly mentioned
[111]	Kinect camera	Not mentioned	LSD-SLAM for assessing the user's location and building an environmental map	•
[112]	Laser Range Finder (LRF) sensor	PC	Hector SLAM for map building and odometry calculation	•

TABLE 20. Classification of sensor type

Sensor Type	References
Camera only	[56], [60], [65], [73], [75], [85], [93], [101], [102], [105], [109]
Camera, IMU	[98]
RGB-D camera	[57], [62], [64], [66], [69], [71], [72], [89], [90], [97]
RGB-D, LiDAR	[100]
RGB-D, IMU	[70], [74], [80], [84], [88], [91], [92], [96], [99]
LiDAR	[61], [67], [68], [112]
LiDAR, IMU	[87], [95]
Stereo camera, other sensors	[63], [104], [108]
3D time-of-flight	[77]
3D time-of-flight, IMU	[78], [106]
Fisheye, depth camera	[76]
Others	[58], [59], [79], [82], [83], [103], [107], [110], [111]

need for lightweight, portable devices with a long battery life is critical for BVI users, who rely on these systems for extended periods. In addition, robust performance in diverse environments, including crowded spaces and varying lighting conditions, is vital for effective BVI navigation. Unique to BVI applications is the need to translate complex spatial data into intuitive, non-visual feedback. Furthermore, the integration of semantic information to provide context about the environment (e.g., identifying doors, stairs, or pedestrian crossings) in order to interact with the environments during navigation is particularly valuable for BVI users, but may be less critical in other SLAM applications. In the following subsections, we explore in detail the advantages and limitations of SLAM techniques when applied to BVI navigation.

1) Advantages

Unlike many localization approaches, such as RFID- or GPS-based methods, which require infrastructure setup, SLAM does not depend on pre-existing infrastructure. It operates autonomously by creating maps and understanding the surroundings in real-time. SLAM relies on data captured by sensors already present on many mobile devices, such as smartphones, and offers a cost-effective solution for accurate localization and mapping without the need for additional hardware or subscription services.

One of the most important advantages of SLAM is its potential for real-time positioning, which determines the agent's current location and orientation. Systems leveraging ORB-SLAM, for instance, excel in pose estimation by integrating various data types, including visual, inertial, and depth information, thus enhancing accuracy beyond conventional methods [64]. This feature is pivotal not only for effective navigation but also for obstacle detection and avoidance, ensuring the safety and confidence of visually impaired users as they navigate through immediate environments [64], [89].

SLAM's ability to reuse and update maps incrementally allows for a high degree of environmental adaptation. Its capacity to relocalize within prebuilt maps or expand them as necessary ensures that users can rely on updated information for navigation [75], [76]. This adaptability is further enhanced by the capability of the system to handle dynamic environments, making it invaluable for visually impaired users who require real-time path adjustments in response to moving obstacles [73], [90].

Detailed environmental mapping facilitated by SLAM, ranging from two-dimensional layouts to complex 3D geometric and semantic maps, provides comprehensive spatial understanding [59], [67], [72]. Environmental awareness is critical for path planning and collision avoidance. Furthermore, the integration of semantic mapping enriches spatial understanding by adding contextually rich information to maps, thereby facilitating more informed decision-making and interaction with the environment [66], [100].

The integration of different types of sensors and technologies with SLAM significantly expands the scope of its application. By integrating techniques such as object detection algorithms or combining RGB-D and IMU sensor data, SLAM systems achieve a multilayered perception of the environment [64], [84]. Sensor fusion enhances a system's ability to detect and classify objects, accurately navigate, and handle dynamic elements within the environment, thereby offering a more holistic assistive solution [84].

The cost-effectiveness of SLAM-based solutions attributed to their reliance on widely available low-cost sensors makes this technology particularly interesting. Systems employing monocular cameras or wearable RGB-D cameras exemplify how SLAM can be implemented in a cost-effective manner without compromising functionality, thus making advanced navigational aids accessible to more users [69], [85]. The advantages of SLAM, derived from the literature, are listed

TABLE 21. Categorization of computing resources used for localization and mapping tasks.

Computing resource	Description	References
Remote/Cloud servers	Used for high-computation tasks, leveraging the power of remote resources.	[57], [60], [63], [66], [72], [92], [93], [108]
Smartphones/Tablets	Common in applications prioritizing portability and accessibility.	[56], [74], [75], [80], [91], [98], [107], [110]
Nvidia Jetson	For balancing computational power and portability.	[60], [62], [71], [82], [84], [85], [87], [97], [103], [104]
Raspberry Pi	Chosen for its affordability and sufficient power.	[64], [67], [69], [79], [83], [89], [104]
Laptops and PCs	Employed for tasks needing robust computational capabilities and flexibility in hardware.	[59], [77], [88], [90], [95], [100], [112]
UP Board computer	Used for handling intensive computation tasks while maintaining a compact form factor.	[70], [78], [99], [106]
Other specific systems	Includes a variety of specific embedded solutions tailored to the requirements of each study.	[58], [68], [72], [76], [89], [102], [104], [109]

in Table 22.

2) Limitations

Although SLAM technologies show great potential for improving navigation aids for the visually impaired, they are not without their limitations. These limitations can significantly impact the effectiveness and reliability of SLAM-based assistive systems.

A notable challenge is the computational complexity and the associated demand for system resources. The implementation of advanced SLAM algorithms and the integration of deep learning frameworks for semantic understanding introduce significant computational overhead [57], [72]. This complexity can compromise the real-time performance, which is crucial for assistive navigation. The need for appropriate hardware to process high-resolution data further underscores this limitation, potentially restricting the deployment of SLAM-based systems [89].

The effectiveness of SLAM is dependent on its environmental characteristics. Accurate mapping and localization depend on the presence of distinct geometric features. In environments lacking such features or dynamically changing settings, SLAM systems may struggle to maintain accurate localization, thereby leading to navigation errors [77]. This limitation is particularly evident in feature-poor areas such as long corridors or spaces with uniform surfaces, where loss of localization can occur [62].

Another critical limitation is the dependence on initial data or pre-existing maps. Some SLAM systems require sighted individuals to pre-map the environment, which can limit the flexibility and immediate usability of unmapped or altered spaces [75], [76]. This reliance on prior mapping can be a significant hurdle for deploying SLAM-based navigation aids in diverse and changing environments [82], [96].

Drifting errors present a substantial challenge for maintaining the long-term accuracy of SLAM systems. Over time, small inaccuracies can accumulate, leading to significant deviations from the true trajectory, which can disorient users and compromise the navigation safety [56]. In addition, the ineffectiveness of some SLAM systems for generating dense navigation maps limits their utility in providing the detailed guidance required for visually impaired navigation, necessitating further algorithmic enhancements [64].

The performance of SLAM in dynamic environments, characterized by moving obstacles and changing conditions, remains a critical concern. Systems may fail to adapt quickly to such changes, leading to potential navigation errors and safety risks for visually impaired users [68].

Some SLAM applications require external calibration or setup, such as placement of calibration boards in specific environments. This requirement can limit the spontaneity and ease of use of SLAM-based navigation aids because it imposes additional constraints [105]. Table 23 outlines the overall limitations of SLAM, derived from the publications under review.

C. RQ3. WHAT CHALLENGING SITUATIONS HAVE BEEN ADDRESSED?

This section explores various challenging situations addressed by SLAM-based navigation-assistive systems for BVI individuals. We categorized these challenges into two main groups: those relevant to environmental complexities and those related to the sensors used for receiving environmental data. Additionally, we discuss practical challenges and considerations that impact the usability and adoption of these systems.

1) Technical and methodological challenges

Optimal pathfinding, perception of surroundings, and obstacle avoidance are crucial for navigation. Precise localization of a visually impaired user within the environment is essential for the effective operation of these functions. Our surveyed papers addressed the localization and mapping problems using various techniques. Some of these studies explored other navigation-related challenges. We categorized these challenges into two groups: those relevant to environmental complexities, and those related to the sensors used to receive environmental data. Dynamic obstacles and crowded spaces constitute the challenges in the first group, whereas challenges related to changing lighting conditions and the rapid motion of users that results in motion blur fall into the second group. Table 24 lists studies that investigated these challenges through the integration of SLAM with other approaches. In the following section, we discuss the studies that address these challenges.

a: Environmental complexities

Crowded scenarios Navigating crowded environments presents significant challenges for the visually impaired, leading to increased collision risks and difficulties in maintaining

TABLE 22. Key features of SLAM that benefit visually impaired navigation.

Advantages	Description	Ref.
Accurate localization	SLAM provides precise positioning within an environment, essential for effective navigation and obstacle avoidance.	[56]–[58], [62], [64], [67]–[70], [73]–[78], [82], [83], [85], [88]–[93], [95], [97]–[99], [102], [106]–[109]
Environmental mapping	SLAM constructs maps of surroundings, enabling spatial awareness for navigation.	[57]–[61], [64], [67], [69], [70], [72]–[77], [79], [83], [84], [87]–[90], [92], [93], [96], [97], [99], [100], [104], [107]–[112]
Map reuse	Previously created maps can be reused to enhance the efficiency and reduce the need for constant remapping.	[57], [75], [76], [93], [109]
Loop closing	SLAM corrects trajectory drifts by recognizing previously visited locations, thereby improving long-term accuracy.	[57], [63], [84]
Semantic mapping	SLAM integrates contextual information into maps, thereby enhancing the understanding of the environment and its elements.	[66], [84], [100]
Object localization	SLAM identifies and positions objects within the environment, facilitating interaction and navigation around the obstacles.	[72], [85], [88], [91]
Dense navigation maps	SLAM generates detailed maps that provide rich environmental data that are crucial for complex navigation tasks.	[64], [66], [89]
Incremental map updating	SLAM continuously updates maps with new information, ensuring that they remain accurate and current.	[64], [69], [89], [109], [111]
Integration with other technologies	SLAM can be combined with other technologies and algorithms to enhance functionality and performance.	[64], [68], [72], [83]–[85], [89], [92], [98], [102]
Integration of sensors	SLAM utilizes a variety of sensors to enrich the environmental perception and mapping accuracy.	[58], [84], [91], [99], [100], [104]
Cost-effectiveness and accessibility	SLAM's reliance on commonly available sensors makes it an affordable solution for widespread use.	[57], [60], [66], [67], [69], [73]–[76], [80], [83], [85], [98], [102]
Ground-truth trajectory	SLAM delivers accurate path tracking, aiding the development of reliable navigation instructions.	[63], [65]
Integration with 2D map	SLAM data can be integrated with 2D maps to enhance the navigation accuracy and functionality.	[63], [70], [75], [77], [80], [98]–[100]
Dynamic environment handling	SLAM can be adapted to changes within environment, maintaining reliable navigation in the presence of moving obstacles.	[73], [90], [100], [104], [109], [111]
Re-localization	SLAM can quickly regain accurate positioning after temporary tracking loss, ensuring continuous and reliable navigation.	[57]

personal spaces. The absence of visual information makes it difficult to measure distance, perceive crowd density, and locate landmarks or places of interest.

Navigation in crowded environments also poses challenges for assistive technologies. For example, in assistive systems that operate based on SLAM, the presence of numerous dynamic elements, such as moving individuals and objects, introduces ambiguity into feature detection and tracking, leading to difficulties in accurately estimating the pose of the user and structure of the environment. The dynamic nature of crowds also hinders loop closure detection, disrupts map consistency, and contributes to drift. Moreover, the lack of distinct visual landmarks in crowded scenes represents a reliable localization challenge, which potentially reduces the robustness and accuracy of the overall SLAM system. Addressing these challenges requires the development of new approaches to address the complexity of such environments effectively. Several studies have investigated this issue.

[59] presented a guide mobile robot engineered for the complexities of navigating different environments while considering dynamic objects and human presence. The robot could handle crowded environments with multiple dynamic objects. To accomplish this, the robot leveraged a spatial risk map, which is a tool that evaluates potential object-occupied spaces, to chart a path that effectively minimizes disruptions. This study presents experiments in which a robot successfully guided a user through the passage of multiple objects and people. The research used Cartography SLAM for off-line mapping. It's important to note that this paper did not address the dynamic environment through SLAM, but rather used SLAM solely as a tool to pre-build the map of the environment.

In another study, [65] introduced an egocentric human trajectory forecasting model that was designed for navigation in crowded environments. The model predicts the path of the sensor wearer using their past trajectories, nearby pedestrian trajectories, scene semantic and depth data. The authors collected an egocentric human trajectory forecasting dataset. As they could not use GPS or motion capture systems for recording the trajectory, they used ORB-SLAM3 to obtain the ground-truth sensor wearer trajectory. The trajectories obtained using ORB-SLAM3 were used to train the egocentric human trajectory forecasting model. It is important to note that this study, like the previous one, did not handle the

TABLE 23. Limitations of SLAM for visually impaired navigation.

Limitations	Description	Ref.
Complexity and computational requirements	SLAM's advanced algorithms can demand significant computational power, impacting real-time performance and efficiency.	[57], [72], [89], [90]
Dependency on environmental features	The accuracy of SLAM depends on the presence of distinct environmental features, which limits its effectiveness in feature-poor or dynamically changing environments.	[77], [109], [111]
Dependence on initial data/prior maps	Some SLAM systems require pre-mapped environments or initial data setups by sighted individuals, thereby reducing the flexibility in unmapped or altered spaces.	[75], [76], [82], [84], [93], [96], [109], [112]
Loss of localization in feature-poor areas	SLAM may experience frequent localization losses in areas lacking sufficient feature points, such as blank corridors or plain walls, thereby compromising navigation reliability.	[62]
Dependency on external calibration	The need for external calibration in certain SLAM applications can limit their spontaneity and practicality in unprepared environments.	[105]
Drifting error	Accumulating drifting errors in SLAM can reduce the long-term accuracy, leading to potential navigation inaccuracies and user disorientation.	[56]
Dense maps fail to align with real-world conditions	The inability of certain SLAM systems to generate detailed dense maps can restrict their effectiveness in providing a comprehensive navigation guidance.	[64]
Vulnerability in dynamic environments	SLAM systems may struggle to adapt to dynamic environments with moving obstacles, thereby posing navigation challenges and safety risks.	[68]

TABLE 24. Challenges addressed in reviewed studies using SLAM techniques

Related to	Challenges	Reference(s)
Environment	Crowded places	[59], [65], [68], [95], [104]
	Dynamic object	[59], [68], [71], [73], [90], [100], [109], [111]
Sensor	Fast motion	[72]
	Illumination	[61], [71], [72], [74], [81]

crowded environment through SLAM itself, but rather used SLAM as a tool for obtaining the trajectory ground-truth.

In addition, [68] addressed challenges in crowded environments using a combination of SLAM and Ultra-Wideband (UWB) positioning. However, the SLAM algorithm was found to be less effective in environments with dynamic obstacles such as pedestrians. The algorithm finds features of dynamic obstacles moving along with the robot as the assistive device, and thus, it was misled that the robot did not move at all. However, using UWB positioning mitigates this issue.

In [95], the method addressed crowded environments by recognizing and predicting people's behavior while anticipating the collision risk. The system advises users to adjust their walking speed (on-path mode) or to choose alternative routes (off-path mode). This involves comparing the 3D point cloud map to real-time LiDAR and IMU sensor data. The system then predicts the future position and velocity of the user in order to avoid collisions.

An intelligent autonomous scooter was developed in [104] for navigating environments with small safety margins and highly dynamic pedestrian traffic such as sidewalks with numerous obstacles and pedestrians. The authors proposed a hybrid mapping solution that combines far-field and near-field mapping to navigate through dynamic environments. This approach utilizes sensor fusion to adapt dynamically to complex and cluttered environments. However, it should be noted that this study conducted system tests in a completely static environment without moving objects, such as pedestrians. Furthermore, the RTAB-MAP SLAM system was used in this study without any adaptation to dynamic environments.

Dynamic objects Several publications have not directly addressed the challenge of crowded environments, and have only focused on dealing with the presence of dynamic objects within the scene. In the system proposed by [90], dynamic objects can be identified, and average depth information can be provided to the user. When a dynamic object belongs to a predefined class such as a person, it can also be tracked between frames in the SLAM pipeline. The system is capable of identifying and tracking dynamic objects after ego-motion estimation to obtain average depth information. Subsequently, it can estimate the poses and speeds of these tracked dynamic objects and relay this information to the users through acoustic feedback. Depth information helps users maintain social distancing in public indoor environments such as shopping malls.

To address the challenge of dynamic objects, [73] proposed a new method called visual simultaneous localization and mapping for moving person tracking (VSLAMMPT). This method was designed to handle dynamic environments in which objects constantly move. The system also uses expected error reduction with active-semi-supervised learning (EER–ASSL)-based person detection to eliminate noisy samples in dynamic environments. This aids in accurate detection and avoidance of dynamic obstacles.

[100] utilized YOLOv3 to detect common objects in a corridor, including people, which were identified as obstacles. The system sends information about obstacles to users every five seconds when the distance between the user and obstacle is less than 10 m. For example, it may notify the user, "A person is located 2.8 meters ahead."

b: Sensor-related challenges

Changes in lighting condition Lighting changes pose a hurdle to visual SLAM systems. Illumination variations alter the visual features, interfere with accurate detection and matching across frames, impact pose estimation, and map-building robustness. SLAM relies on distinctive features for operation; however, lighting changes introduce ambiguities, noise, and errors, which affect the accuracy. Overcoming this challenge requires robust algorithms for dynamic lighting to ensure stable and precise localization and mapping.

The method proposed in [71] tackles the challenge of illumination changes using a deep descriptor network called a Dual Desc, which is designed to be robust against various appearance variations including illuminance changes. The network used multimodal images (RGB, Infrared, and Depth) to generate robust attentive global descriptors and local features. These descriptors were used to retrieve coarse candidates from query images, and 2D local features, along with a 3D sparse point cloud, were used for geometric verification to select the optimal results from the retrieved candidates. The authors mentioned that their dataset included images captured at different times of the day, which resulted in illumination changes between the query and database images. Despite these changes, the proposed method achieved satisfactory localization results.

The authors of [72] evaluated the influence of lighting conditions on the performance of their novel localization method. The authors captured training images during the day and test images at night and simulated changes in lighting conditions by switching some of the lights off in locations without windows. The results showed that changes in the lighting conditions had a minor impact on the proposed method.

[61] mentioned that the proposed localization scheme was verified in a typical office building environment with dramatically changing lighting conditions throughout the day; however, it does not provide detailed results or discussion on how changing lighting conditions affect the performance of the system.

The method proposed by [81] addressed changes in illumination as a challenge using the COLD and IDOL datasets, which were recorded under different weather and illumination conditions (cloudy, night, and sunny) using different mobile platforms and camera setups. These datasets were used to evaluate the strength of the localization and recognition algorithms with respect to the variations caused by human activities and changes in illumination conditions. The study also mentioned the use of Histogram of Oriented Gradients (HOG) for feature extraction, which provides preferable invariant results for lighting and shadowing.

Motion blur Blurred images in visual SLAM can lead to inaccuracies in feature detection and matching, causing issues with pose estimation, map building, loop closure, and visual odometry. These inaccuracies can also impact depth measurements and map quality. To address this issue, strategies such as using high-frame-rate sensors, incorporating IMUs, and employing motion-deblurring techniques can be employed to improve the accuracy of localization and mapping in SLAM systems.

Motion blur can be caused by fast or sudden movements of the user during navigation, which can affect localization performance. [72] studied this challenge and evaluated the robustness of the localization methods. They captured 2316 blurred images on the testing day. The results show that the proposed method performed poorly in this experiment, indicating that fast motion or sudden changes in user movement can pose a challenge to the system. The reason for this poor performance is that the object detection scores did not exceed the threshold during the experiment.

2) Practical challenges and considerations

In addition to the technical and methodological aspects, we recognized the importance of practical challenges and considerations that can significantly affect the usability and adoption of SLAM-based assistive systems. Therefore, we included an evaluation of the practical challenges and operational efficiency, as summarized in Tables 25 and 26. The information in these tables has either been directly extracted from the article or can be easily inferred from the article's text. These tables provide information on user-friendliness, cost-efficiency, weight, comfort for extended use, adjustable fit, fatigue mitigation, and portability of the assistive tools described in the reviewed studies. For instance, while smartphones and lightweight devices such as eyeglasses-mounted sensors [74] and ARCore-supported smartphones with haptic gloves [75] are generally well received because of their high portability and ease of use, heavier devices such as guiding robots [68] and rolling suitcase-shaped device [95] are noted to cause user fatigue over extended periods. The augmented cane [67], although found to improve confidence and workload for novice and expert users, also faced usability challenges owing to its weight. Clear instructions and easy learning curves, as seen in electronic glasses with haptic modules [57], play a significant role in enhancing user satisfaction. However, the cost efficiency of these technologies varies, with some solutions being more affordable and accessible.

D. RQ4. HOW THE PROPOSED SOLUTION IS EXPECTED TO ENHANCE MOBILITY AND NAVIGATION FOR VISUALLY IMPAIRED?

This section discusses how the approaches proposed by the studies included in our SLR have the potential to improve navigation for BVI people. These studies focused on diverse attributes, such as accurate pose estimation, semantic mapping, sensor fusion, and algorithmic innovations, to improve the quality of BVI navigation. Table 27 presents the categorization of attributes that contribute to enhancing the mobility and navigation of visually impaired individuals.

To understand the impact of these solutions further, we examined their effectiveness in real-world scenarios. Various localization and mapping techniques have been assessed on the basis of their accuracy, robustness, consideration of dynamic objects, and running time. This evaluation provides insight into the performance of these techniques in practical environments.

In addition, we considered user-based evaluations to gauge user satisfaction and the practical applicability of the proposed system. These evaluations include feedback from actual users, which is crucial for understanding real-world usability and acceptance of assistive technologies.

TABLE 25. Practical challenges and operational efficiency - Part I.

Ref.	Assistive tool	User-friendliness		Cost-efficient	Weight	Comfort for extended use	Adjustable fit	Fatigue mitigation	Portability
		Clear instructions	Easy to learn						
[56]	Smartphone	✓	✓						
[57]	Electronic glasses and leg-mounted haptic modules	✓		✓	Light	✓	✓	✓	High
[58]	Microsoft Hololens2	✓			Light		✓		High
[59]	Wheeled guide mobile robot	✓			Light	✓ Easy-to-hold handle			Moderate
[60]	Android application	✓		✓	Light	✓	✓	✓	High
[61]		Not a user-based evaluation: only a technical test.							
[62]	Helmet; white cane	✓			Light	✓			Moderate
[63]	Smartphone	✓	Need time to learn interpreting tactile signals						
[64]	Smart cane	✓			Light				High
[65]		No prototype implemented; the proposed approach for trajectory forecasting was tested with a robot.							
[66]	A forehead-mounted camera, an earphone, a computing resource bag	Not a user-based evaluation: only a technical test performed.		✓	Heavy	Heavyweight		Heavyweight	Moderate
[67]	Augmented cane	✓	✓	✓	Heavy (1kg)	Heavyweight	✓	Heavyweight	Moderate
[68]	Guiding robot	✓	✓	USD 17000	Heavy (25kg)	Comfortable handle feedback; some users noted slewing and speed change discomfort.			The robot is relatively large at 41x43x25 cm^3.
[69]	Head-mounted camera	Not a user-based evaluation: only a technical test performed.		✓	Light				
[70]	Smart cane	Not a user-based evaluation: only a technical test performed.							
[71]	Auxiliary glasses	Not a user-based evaluation: only a technical test performed.							
[72]		No prototype implemented: Only a technical test performed.							
[73]	Smart eyeglasses	Limited instructions		✓					High
[74]	Eyeglasses-mounted sensors + smartphone	✓	Participants trained in 10 minutes	✓	Light	Prolonged beeping may cause discomfort.		Glasses weight strains the nose.	High
[75]	ARCore-supported smartphone + haptic gloves	✓	A 5-minute tutorial	✓					High
[76]	Optical see-through glasses			✓					High
[77]	Smart cane				Heavy	Low comfort		Heavyweight	Bulky tablet hangs on neck, heavy cane
[78]	Computer-vision-enhanced white cane	Not a user-based evaluation: only a technical test performed.							
[79]	Smart robot	No information available.							
[80]	Smartphone			✓	✓	✓	✓	✓	High

Furthermore, we provide a detailed overview of the components and technologies used in assistive navigation systems. This helps to understand the practical implementations and innovations proposed by the studies. By examining the system prototypes, we gained insights into the design and functionality of assistive solutions beyond the localization and mapping components. This offers a comprehensive view of how these technologies enhance the mobility and navigation of the visually impaired.

1) Attributes enhancing mobility and navigation

SLAM technology is primarily used to provide precise localization, which is critical for assistive navigation systems. Precise localization provides accurate information regarding a user's position in the environment in which the user navigates. This accuracy enables the system to offer feedback on obstacles, pathways, and points of interest, thereby allowing BVI to navigate safely and confidently. Real-time assistance has also emerged as the key feature. Providing immediate feedback on the environment enables users to travel efficiently and safely, which leads to increased mobility and independence. Semantic mapping generates maps beyond geometric data.

TABLE 26. Practical challenges and operational efficiency - Part II.

Ref.	Assistive tool	User-friendliness		Cost-efficient	Weight	Comfort for extended use	Adjustable fit	Fatigue mitigation	Portability
		Clear instructions	Easy to learn						
[82]	Head-worn camera	Easily understood commands	10-minute training	Commercially available hardware	Heavy (5.5lbs ≈ 2.49kg)				High
[83]	Sensors attached to a white cane	Not a user-based evaluation.		✓	Not a user-based evaluation: only a technical test performed.				
[84]		No prototype implemented; the proposed approach for real-time global localization was tested with an agent.							
[85]	Chest-mounted camera	Not a user-based evaluation		✓	Light	Not explicitly mentioned.			High
[87]	Suitcase-shaped robot	Improved with intuitive terminology	Positive usability scores	-	Heavy (40lbs ≈ 18.14 kg)	Physical demand	-	Heavyweight	Bulky and heavy
[88]	Robotic cane								
[89]	Smart cane	The focus of the paper is primarily on the technical aspects of the multi-sensory blind guidance system.							
[90]	Smart glasses				Light				
[91]	Hand-worn device	Not a user-based evaluation: only a technical test performed.			52.5 grams				
[92]	Smart E-glasses					Areas for improvement			
[93]	Google glasses	Information not provided; Only a technical test conducted with a laptop PC as a navigator.							
[95]	A rolling suitcase-shaped device	✓	Short training session (10-20 minutes)		Heavy	Weight discomfort		Heavy and bulky	Space limitations
[96]		System testing was not feasible as the system was in the experimental stage.							
[97]	Smart glasses	Not a user-based evaluation: only a technical test performed.							
[98]	Smartphone	✓	✓	✓	Light	✓	✓	✓	High
[99]	White-cane mounted camera	Information not provided; Only a technical test performed.							
[100]	TurtleBot2 robot (to be replaced by portable device) + wearable sensors				Heavy	Heavy and bulky			Bulky
[101]	Paper develops a Reinforcement Learning environment to create a navigation assistant tailored for the BVI community, without user-based evaluation.								
[102]	Smartphone			✓	Light	✓	✓	✓	High
[103]		No prototype implemented; the proposed approach was tested within a research building.							
[104]	Intelligent autonomous scooter			High cost	Heavy				Bulky
[105]	Wearable camera				Light				High
[106]	Computer-vision-enhanced white cane	Not a user-based evaluation: only a technical test performed.							
[107]	Smart cane + Google Tango				225 grams				High: light and compact, usable in multi-floor
[108]	Helmet-mounted camera + android-based smartphone	Not a user-based evaluation: only a technical test performed.							
[109]	Wearable camera	Not a user-based evaluation: only a technical test performed.							
[110]	iPhone 12 Pro Max	Effective language instructions	Intuitive navigation process		Light			-	High
[111]	Wearable camera	Not a user-based evaluation: only a technical test performed.							
[112]	Person carrier robot (wheelchair)				Heavy				Bulky

Such representations contain not only spatial information but also the semantic meanings of objects and features within the environment. This semantic understanding offers a deeper insight into the environment. This contextual awareness is particularly beneficial for enhancing navigation accuracy, as it enables navigation systems to make decisions based on semantic context, improving obstacle avoidance, path planning, and overall navigation efficiency.

By employing robotic systems such as small robots, smart canes, and sensor-equipped suitcases, some studies have provided guidance, obstacle avoidance capabilities, and increased spatial awareness, thereby effectively providing independent navigation for visually impaired individuals. Smartphone-based solutions harness the ubiquitous nature of smartphones that are equipped with cameras and sensors. These solutions offer navigation assistance by using widely available and familiar devices. Both indoor and outdoor navigation capabilities offer a seamless transition between environments, ensuring that users receive consistent support, regardless of the scene in which they navigate. Innovative localization and mapping algorithms enhance navigation efficiency and effectiveness through tailored modifications of existing SLAM frameworks or through the creation of novel solutions. Ultimately, these advancements have led to an improved overall experience for individuals with visual impairment. Although these studies focused on different features and attributes, they all aimed to enhance mobility, independence, and overall quality of life for BVI people.

2) Effectiveness of localization and mapping techniques in real-world scenarios

The effectiveness of the localization and mapping techniques in real-world scenarios varies across studies. Tables 28-30 summarize these evaluations, highlighting key attributes such as the working area, localization and mapping accuracy level, robustness level, consideration of dynamic objects, and running time. The robustness and accuracy levels reported in these tables are extracted from each paper's context; each rating reflects conditions specific to that paper and is not necessarily superior or inferior to the other approaches. Thus, these values are not comparable due to differing conditions across the papers.

Many studies, such as [57], [59], and [62], have demonstrated high localization and mapping accuracy, particularly in indoor environments. These studies employed techniques such as ORB-SLAM2 and Cartographer to ensure reliable feature matching and adaptive navigation.

Robustness is another critical factor, with many systems proving to be resilient under various conditions. Studies such as [61] and [71] reported high robustness owing to the integration of multiple sensors and multi-modal imaging. These systems can navigate complex environments and maintain accurate localization.

However, some studies highlighted some challenges. For instance, [68] indicated that SLAM-based systems struggle with dynamic environments, leading to unstable navigation and orientation errors. Similarly, [82]pointed out issues with SLAM-relative poses in changing or occluded feature scenarios that affect navigation stability.

The running time is another essential consideration, with many studies emphasizing real-time performance. Systems such as those described in [66] and [67] provide real-time performance, which is crucial for assistive navigation. However, some systems, such as those in [72], face longer computational times owing to their increased complexity, which can be a drawback in real-world applications.

Overall, the evaluation of localization and mapping techniques across different studies revealed a range of performance levels. High accuracy and robustness are common in controlled indoor environments, whereas dynamic and complex scenarios pose significant challenges. Insights from these evaluations are crucial for understanding the practical applicability and limitations of SLAM-based assistive systems for visually impaired individuals.

3) User-based evaluations

This section analyzes the user-based evaluations conducted to assess the satisfaction of the proposed SLAM-based assistive systems. By examining these evaluations, we gained insights into the real-world applicability and user acceptance of these technologies. Several studies conducted user-based evaluations with actual participants to assess the effectiveness of and satisfaction with their proposed systems. These evaluations provided valuable insights into the usability and acceptance of assistive technologies. Tables 31 and 32 summarize studies that include user-based evaluations.

Three methods for assessing user satisfaction were identified: user studies, interviews, and surveys. Additionally, some studies involved only visually impaired participants, some involved only blindfolded users, and some included both groups to test their systems. Most studies used user studies as the primary evaluation method. Some studies also employed interviews or surveys after initial user studies to gather additional information on user satisfaction. The tables also show the experimental sites where the evaluations were conducted.

For example, [56] involved nine BVI participants on a university campus to evaluate a sonification system and collect feedback on pleasantness, annoyance, precision, quickness, and overall appreciation. Similarly, [57] conducted evaluations with two BVI and three blindfolded participants in a laboratory setting, focusing on the task success rates, completion times, and feedback from verbal and haptic cues.

The study by [58] included five BVI and three blindfolded participants, achieving user satisfaction scores between six and nine out of ten. Another study by [59] evaluated their system with ten blindfolded participants, noting improvements in acceptance and trust levels.

[62] found moderate to high satisfaction among eight blindfolded participants, who found the system acceptable and useful for indoor navigation. In contrast, [77] highlighted that while users found the wayfinding function useful, they expressed discomfort owing to the weight of the device. Overall, the user-based evaluations indicated that participants generally found the proposed systems beneficial for navigation, with varying levels of satisfaction based on the specific features and implementation of each system.

Some studies only conducted technical tests, without involving direct user feedback. These studies are summarized in Table 33. For example, [65] and [66] focused on the technical performance of their systems and conducted tests in controlled environments but did not report user satisfaction.

The absence of user-based evaluations limits our understanding of how these systems perform in real-world scenarios and their acceptance among users. Future research should aim to incorporate comprehensive user studies to complement technical assessments and provide a more holistic view of a system's effectiveness and usability.

4) System prototype information

To provide a comprehensive understanding of the assistive solutions proposed in the reviewed studies, we present the information regarding the system prototypes in Tables 34-38. These tables include data on the functionalities, sensors used, computing resources, human-computer interaction (HCI) mechanisms, assistive tools, battery life, and whether

TABLE 27. Attributes of SLAM-based navigation systems that contribute to enhancing BVI navigation, along with referenced papers emphasizing each feature.

Features	Description	Reference(s)
Precise Localization	Using SLAM algorithms, these solutions accurately estimate the position and orientation of visually impaired users. Precise localization is essential for visually impaired navigation systems in order to ensure accurate real-time guidance, obstacle avoidance, and spatial awareness, ultimately enhancing independent and safe mobility.	[64], [67], [69], [70], [73], [75]–[78], [82], [89], [90], [93], [95], [97], [102]
Real-time assistance	This feature ensures that BVI users receive immediate feedback about their environment. Therefore, they can be guided to navigate safely and efficiently, thereby enhancing their overall mobility and independence.	[57], [66], [67], [69], [77], [82], [85], [88], [92]
Semantic mapping	Semantic mapping involves creating detailed environmental representations that go beyond geometric data and enables visually impaired users to navigate with a deeper understanding of their surroundings.	[66], [72], [84], [97], [100]
Both indoor and outdoor navigation	Systems that serve both indoor and outdoor settings allow users to transition seamlessly between different environments while receiving consistent support.	[56], [65], [67], [68], [74], [85], [104], [108], [111]
Innovative algorithms	Innovative algorithms lead to advancements in navigation techniques. These approaches contribute to a more efficient and effective navigation, ultimately improving the overall experience of the visually impaired.	[66], [73], [77], [78], [91], [102], [106], [110]
Robotic navigation	These solutions employ robotic systems, such as robots, smart canes, scooter, and suitcases, which are equipped with sensors to assist BVI in navigating their environment.	[59], [64], [67], [68], [70], [78], [87], [89], [99], [104], [106], [107], [112]
Smartphone based	Equipped with cameras and sensors, tablets and smartphones can be considered versatile navigation tools. These solutions offer assistance through devices that are widely available and familiar.	[63], [74], [75], [80], [98], [102], [104], [107]–[111]

the solutions are machine learning-based. Notably, the specifications in these tables cover the entire assistive system, and not just the localization and mapping components, as presented in Tables 17-19.

Functionalities include the capabilities and features of the assistive system, such as navigation, object recognition, and obstacle avoidance. Sensors specify the types of sensors used in assistive devices such as cameras, LiDAR, and IMUs. Computing Resource indicates the hardware used for processing, including local devices, such as smartphones and laptops, as well as remote servers. HCI describes the interaction mechanisms used to provide feedback to the user, such as audio and haptic feedback. The assistive tool details the form factor of assistive devices, such as smart glasses, canes, and robot systems. Battery Life provides information on the operational duration of a device on a single charge. ML-based indicates whether the assistive solution incorporates machine learning algorithms.

Table 39 categorizes the papers based on the functionalities offered by assistive systems, highlighting the diverse capabilities ranging from basic navigation and obstacle avoidance to advanced features such as scene understanding and social networking. Table 40 classifies the papers based on the HCI mechanisms employed, showing the prevalence of audio feedback and the growing trend towards multimodal feedback incorporating haptic and tactile cues. Finally, Table 41 categorizes the studies based on the form factor of the assistive tool, revealing the diversity of approaches, including smartphone-based, wearable devices, handheld devices, and robotic systems.

a: Functionalities

The analysis of the data in the tables shows the diverse range of approaches and technologies used to create assistive systems for visually impaired navigation. Most systems focus on navigation and obstacle avoidance, but many also include advanced features such as scene understanding and social networking. The use of sensors is diverse, with RGB-D cameras being the most commonly used because of their capability to capture both color and depth information, especially for the localization and mapping components of assistive systems.

b: HCI

The HCI mechanisms vary, with audio feedback being the most commonly used method. Several systems also use haptic feedback and a few incorporate visual hints for users with partial vision. These feedback mechanisms are essential for real-time navigation assistance and for enhancing user experience. Several studies used multimodal feedback for real-time navigation assistance, as indicated by the HCI column. These include combinations of audio, haptic, and grounded kinesthetics to enhance user experience and provide comprehensive navigation aids. This multimodal approach ensures that users receive complementary information, thereby enhancing the robustness and reliability of assistive systems.

c: Assistive tool

The form factors of assistive tools vary among the studies. Wearable devices such as smart glasses and helmets are designed to be worn on the body and provide hands-free assistance. Handheld devices, such as smart canes, are traditional mobility aids enhanced by modern technology. These smart canes include sensors to detect obstacles and provide real-time feedback through vibrotactiles or steering. This approach leverages the familiarity and comfort of using a cane, while adding significant technological advancements to aid navigation and spatial awareness. Some prototypes incorporate both wearable and handheld components; for example,

TABLE 28. **Effectiveness of localization and mapping techniques in indoor environments as indicated by the literature. (Part I)**

Ref.	Localization & mapping accuracy level	Robustness level	Considers dynamic object	Running time
[57]	High: Ensuring safe indoor navigation	High: Utilizes robust ORB-SLAM2 technique for reliable feature matching and adaptive navigation.	•	Affected by 0.15-1 sec transmission time to remote server
[58]	Not provide specific details	High: Instant localization, robust map building	•	Real-time
[59]	High: Confirmed through user-guided tests	Not provide specific details	✓	Not mentioned
[60]	High: With an average error of less than 1 meter	High: Combined advantages of OpenVSLAM and Colmap	•	High: Average response time of 2 to 3 seconds for localization
[61]	High: With 0.62m for 3D and 1.24m for 2D	High: Cartographer for mapping with optimization techniques	•	Moderate: For localization is under 0.25 seconds
[62]	High: Accurate mapping, guiding users precisely to target goals	High: Reliable performance with accurate mapping	•	Moderate: Some delays in mapping, but overall meets real-time requirements
[63]	Moderate: Needs improved SLAM stability and accuracy due to scaling issues	High: due to SLAM loop closing, real-time correction	•	Not mentioned
[64]	Moderate: Effective mapping and localization but has deviations and mismatches impacting overall accuracy	High: The integration of ORB-SLAM with YOLO ensures robust navigation and obstacle detection in various environments	•	Moderate: Real-time map-building but slow path planning
[66]	High: With centimeter-level accuracy	High: Real-time performance, centimeter-level accuracy, semantic mapping integration, and resource optimization	•	Real-time performance ensured through careful computing power allocation
[69]	Not explicitly mentioned	High: Hector SLAM algorithm ensures robust mapping and localization accuracy.	✓	Real-time
[70]	High: With significant error reduction and effective 2D mapping alignment.	High: Enhanced by floor plan integration, error reduction techniques, and superior performance in real-time pose estimation.	•	Real-time
[72]	High: Achieved 96.3% localization accuracy; outperforms other models despite pre-training limitations.	High: Demonstrates robustness through accurate SLAM technique, semantic mapping, and deep learning integration for indoor localization.	•	High: Longer computational time due to increased complexity or resource requirements.
[73]	High: Demonstrates superior accuracy compared to ORB-SLAM2 in dynamic environments.	High: Demonstrates robustness through advanced SLAM techniques for obstacle removal and dynamic environment adaptability.	✓	Not mentioned
[75]	High: Rigorous comparison and real-time reliance on CAD maps ensure precise localization.	High: Advanced localization and flexible path planning for accurate and safe navigation.	•	Not mentioned
[76]	High: Key-frame matching and fisheye camera enhance feature detection and accuracy.	High: ORB-SLAM2 for precise localization and obstacle avoidance in effective indoor navigation.	Dynamic obstacle in experiment path	Short: Visual SLAM and dynamic subgoal selection optimize efficiency for real-time localization and mapping.
[77]	High: Superior to plane-based graph SLAM	High: Efficiently addresses all 6-DOF and outperforms traditional SLAM methods	•	59.4 ms average per frame
[78]	High: Demonstrated by superior performance in pose estimation accuracy and robustness in various environments.	High: Enhanced by integrating plane features and employing a plane consistency check for accurate pose estimation.	•	Not mentioned
[79]	Not mentioned	Not mentioned	✓	Real-time
[80]	High: Utilizes visual landmarks, real-time data analysis, addresses limiting factors	High: Robust to superficial changes, requiring updates only for major structural alterations.	•	Moderate: Impacted by the number and type of visual landmarks.
[83]	High: Accurate localization and mapping demonstrated through real-time indoor experiments in static scenarios.	High: Hector SLAM ensures robustness in complex indoor environments for accurate navigation in static scenarios.	•	Real-time
[84]	High: $<$1m position error, $<5^\circ$orientation error	High: Integration of semantic SLAM, optimization of semantic Point Clouds during loop closures	•	Real-time (typically under 10 seconds, faster near distinctive features)
[87]	High: Cartographer SLAM and 360 LiDAR for precise real-time mapping.	High: Cartographer ensures dynamic, real-time LiDAR mapping with efficient updates in unfamiliar environments.	•	Real-time
[88]	High: SLAM algorithm and low-drift IMU enable precise user pose estimation and environment mapping.	High: With real-time updates and effective threshold management.	•	Short: Navigation completed within 45 seconds on average, well under the 2.5-minute cut-off.
[89]	High: Highly accurate in real-time testing	High: Integrating ORB-SLAM and YOLO ensures stability and accuracy.	•	Not mentioned
[90]	High accuracy feature-based visual SLAM estimation	Moderate: Acknowledging limitations while utilizing effective techniques, indicating room for improvement.	✓	Moderate: Accounting for variable GPU impact and dependency on tracked objects.

in the study by [107], a Google Tango device was mounted on the user's chest while the user held a smart cane. These prototypes are categorized as handheld devices because the users' hands are occupied. Robotic systems represent another innovative factor. These can range from small mobile robots in the shape of a suitcase that guides users through complex environments to more substantial ride-on systems such as autonomous wheelchairs or scooters.

d: Battery life

The battery life is a critical factor in assistive navigation systems. These systems must be reliable to ensure continuous assistance without frequent recharging interruptions. The battery lives of the proposed solutions varied across the reviewed studies. Some systems, such as those described by [57], have reported a long battery life, which ensures that the devices remain functional during extended use. However, not all studies provide detailed information on battery life. This lack of information can be a concern, as it leaves un-

TABLE 29. Effectiveness of localization and mapping techniques in indoor environments as indicated by the literature. NM indicates that the information is not explicitly mentioned in the paper. (Part II)

Ref.	Localization & mapping accuracy level	Robustness level	Considers dynamic object	Running time
[91]	High accuracy (RMSE 0.269), surpassing VINS-Fusion and VINS-RGBD.	High robustness with RGBD-VIO, enhancing accuracy and efficiency for assistive navigation and object manipulation.	•	Not mentioned
[92]	Achieving a 91% success rate in navigation tasks	Demonstrating robustness, utilizing ORB-SLAM2 algorithm enables real-time path planning.	•	Real-time
[93]	Not specified	Not specified	•	Not mentioned
[95]	High: LiDAR and IMU data ensure precise localization and mapping.	High: Integrating multiple sensors ensures robust localization for accurate navigation and collision avoidance.	✓	Not mentioned
[96]	Lacking explicit discussion on accuracy, robustness, consideration of dynamic objects, and ML-based solutions.			
[97]	While numerical metrics are not mentioned, indicators suggest potential accuracy.	High: OpenVSLAM framework ensures robust real-time mapping and localization capabilities.	•	Real-time
[98]	High: Consistently achieves sub-1 meter localization accuracy upon algorithm convergence.	High: Utilizing stable landmarks and a particle filter ensures robust indoor localization.	•	Real-time operation with accurate localization post-algorithm convergence, demonstrating efficient performance on smartphones.
[99]	High: Achieving accurate pose estimation with a position error of 0.2 meters.	High: Enhancing robustness with integrated VIO and Human Intent Detection for accurate pose estimation and mode selection.	•	Short: Real-time pose estimation with updates every 22 milliseconds.
[100]	Moderate Based on validation conducted in real-world scenarios	Low: Performance degradation observed in specific scenarios due to narrow corridor and orientation estimation issues.	✓	Real-time
[102]	High: 94-98% with relative errors of 1.6-2.6%	High: Efficient integration of visual SLAM, object detection, and depth measurements for precise, reliable indoor navigation.	•	Real-time
[103]	Not mentioned	Not mentioned	•	Not mentioned
[106]	High: Superior pose estimation accuracy (mean End Point Error Norm (EPEN): 2.63%) compared to a state-of-the-art VIO (mean EPEN: 6.06%).	High: More stable performance (standard deviation EPEN: 1.3%) than a state-of-the-art VIO (standard deviation: 8.22%).	•	Not mentioned
[107]	Not evaluated	Not evaluated	•	Real-time
[109]	High: Precise self-positioning and mapping with high accuracy in large-scale environments.	High: Superior robustness compared to conventional monocular SLAM algorithms, ensuring quick and reliable calculations.	✓	Real-time
[110]	High: Assessed via quantitative metrics and real-world tests, achieving low navigation error and high success rate.	High: Utilizing robust visual-inertial SLAM with iOS ARKit for accurate real-time navigation and obstacle avoidance.	NM	Real-time per-frame pose estimation
[112]	Not directly mentioned	Moderate: Issues with mapping reflective or transparent surfaces like glass windows.	•	Not mentioned

certainty regarding the reliability of the device in real-world scenarios. Additionally, some devices, such as those incorporating high-performance processors or multiple sensors, may face challenges in maintaining a long battery life owing to their higher power consumption. Systems such as those described by [82], which use advanced components such as the Nvidia Jetson AGX Xavier, may offer robust functionality, but require careful management of power resources to ensure adequate battery life.

e: Machine-learning approaches

Many systems leverage machine learning for functionalities, such as object detection, scene understanding, and localization. Algorithms such as YOLO, Faster R-CNN, and various deep neural networks are commonly employed. These machine learning-based solutions enhance the accuracy and efficiency of the assistive systems. Table 42 categorizes the machine learning approaches used in assistive devices, along with their references. This categorization illustrates the diversity of machine-learning techniques applied to improve the functionalities of assistive systems for BVI navigation.

IV. FUTURE OPPORTUNITIES

An effective navigation system for BVI people needs to meet mobility metrics, such as decreasing navigation time, decreasing navigation distance, decreasing contact with the environment, and increasing walking speed. These systems must be highly accurate and efficient in complex situations, such as crowded places and changing light and weather conditions. At the same time, assistive aids should be comfortable, easy to use, unobtrusive, cost-effective, lightweight, and reduce cognitive load.

In this review, we examined publications that employed SLAM techniques in their navigation approaches. One of the distinct advantages of SLAM is its applicability in diverse locations without the need for pre-built maps or additional infrastructure such as Bluetooth beacons or RFID tags. However, there is still room for improvement in various aspects of these systems, including their ability to handle complex scenarios, provide accurate obstacle information, and seamlessly transition between indoor and outdoor environments.

There is a notable lack of focus on adapting SLAM to challenging situations especially dynamic environments, specifically for visually impaired navigation. Studies in this field typically employ SLAM as a pre-existing tool or use it without any significant adaptation to effectively handle

TABLE 30. Effectiveness of localization and mapping techniques in outdoor and mixed environments as indicated by the literature. NM indicates that the information is not explicitly mentioned in the paper.

Ref.	Localization & mapping accuracy level	Robustness level	Considers dynamic object	Running time
		Outdoor		
[71]	High: Superior accuracy demonstrated across multiple indicators and challenging conditions.	High: Demonstrates superior robustness, especially in the presence of dynamic objects and changing illumination.	✓	Moderate: Meets real-time requirements on powerful devices, but slower on less capable hardware.
[82]	Moderate: Global pose estimation within 80 cm radius of ground truth; verified in real-world scenarios.	Low: Navigation relies on SLAM-relative poses, prone to instability with changing or occluded features.	•	Real-time
[101]	SLAM used for localizing footage, creating spatial graphs, and generating realistic Deep Reinforcement Learning simulator data for pedestrian navigation training.			
[105]	Approximately 12cm error in localization accuracy	High: Robust in weakly textured environments	•	Not mentioned
		Both (Indoor and Outdoor)		
[56]	Moderate: Support in measuring errors.	Moderate: System evaluates errors reliably.	•	Not mentioned
[65]	High: Precise ground truth trajectories, with minimal absolute error	High: Reliable trajectory extraction with low absolute error	✓	Not mentioned
[67]	High: RMSE between 0.08 and 0.44 m, indicating high precision in indoor environments	High: SLAM-based system navigated complex environments with precision and consistent success across multiple trials	•	SLAM operated at 1.4 Hz, which is sufficient for real-time use
[68]	Low: SLAM struggles with dynamic environments	Low: SLAM demonstrates vulnerability in dynamic environments, with unstable navigation and orientation errors compared to UWB positioning	✓	High: Averaging 317 seconds for navigation tasks in a dynamic environment
[74]	High: Indoor localization with pre-built maps.	High: VSLAM for indoor navigation enhances robustness in GPS-degraded environments.	✓	A real-time performance (approximate 20 fps) on a smartphone.
[85]	High: With an error of less than 0.5 meter	High: Integration object detection and visual SLAM for accurate navigation support.	•	Real-time (with initialization under 2 seconds and trajectory estimation under 1 second.)
[104]	High: Accurate obstacle detection, extended range, minimal error in indoor mapping, integrated sensor data for dynamic environments.	High: Dynamic adaptation, effective in complex and crowded environments with moving obstacles.	✓	Not mentioned
[108]	High: Precise image feature extraction and motion trajectory reconstruction	High: Dependable for tracking position and orientation.	NM	Real-time
[111]	Not mentioned	Not mentioned	✓	Not mentioned

challenges. Researchers in this area can draw inspiration from recent advancements in robotics to address this gap. By leveraging cutting-edge techniques from robotic navigation, SLAM systems that are more robust and suitable for assisting visually impaired individuals in real-world scenarios can be developed.

This section discusses the open problems and research directions identified during the SLR.

Challenge scenarios and real-world studies Navigating crowded environments remains a significant challenge for the visually impaired and studies addressing this issue are limited. Evaluations are often conducted in controlled settings rather than real-world scenarios. This is problematic because controlled environments may not accurately reflect the dynamic and unpredictable nature of real-world crowded spaces, where visually impaired individuals face numerous obstacles and safety risks. Future research should focus on developing and testing solutions in high-traffic public places such as train stations and shopping malls. This is crucial to ensure that assistive navigation systems can effectively handle the complexities of real-world crowded environments, including the presence of numerous dynamic objects, varying crowd densities, and unpredictable pedestrian behaviors.

Furthermore, addressing challenging conditions such as changes in illumination, low-light scenarios, high-speed dynamic objects, and complex backgrounds can enhance the robustness and versatility of navigation systems. These challenging conditions are common in real-world scenarios and can significantly impact the performance and reliability of SLAM-based navigation systems. By addressing these challenges, researchers can develop more robust and adaptable solutions that can function effectively in diverse and demanding environments.

Techniques such as image enhancement for ORB points and LSD line feature recovery used in agricultural environments [116] can be adapted for visually impaired navigation. Adapting these techniques from other domains can accelerate the development of more effective solutions for visually impaired navigation, leveraging existing knowledge and expertise to address the specific challenges faced by this user group.

Long-term navigation The development of solutions that are effective over extended navigation periods is critical to achieve autonomous navigation. These solutions must ensure accurate mapping and localization even when the maps are updated over a longer navigation duration. This is important because environments are not static; they change over time. Obstacles may appear or disappear, and landmarks may be altered. A SLAM-based navigation system that cannot adapt to these changes will become increasingly inaccurate and unreliable over time, potentially leading to dangerous situations for visually impaired users.

To address this challenge, researchers can leverage solutions proposed in robotics, such as those presented in [117], which introduced a novel long-term SLAM system with map prediction and dynamic removal, thereby allowing wheelchair robots to maintain precise navigation capabilities over extended periods.

TABLE 31. User satisfaction evaluation. This table includes studies that conducted user-based evaluations with actual participants to assess the effectiveness and satisfaction of the proposed systems - Part I.

Ref.	# Participants		Method of evaluation			Experimental site	User satisfaction
	BVI	Blindfolded	User study	Interview	Survey		
[56]	9	0	✓	•	✓	A university campus	Participants' feedback varied on pleasantness, annoyance, precision, quickness, and overall appreciation of sonification.
[57]	2	3	✓	•	•	Wearable Robotics and Autonomous Unmanned Systems Laboratory at the University of Science and Technology of China	No overall satisfaction score; details on task success rates, completion times, and verbal/haptic feedback.
[58]	5	3	✓	•	•	Not mentioned	Scores 6-9 out of 10
[59]	0	10	✓	•	✓	Not mentioned	Improved acceptance and trust levels noted
[60]	2	4	✓	•	•	New York University Langone Ambulatory Care Center (A Complex hospital environmnet)	Not mentioned
[62]	0	8	✓	✓	•	A room	Moderate to high satisfaction: Users found the system acceptable and useful for indoor navigation
[67]	12	12	✓	•	✓	Hallways constructed with cardboard, outdoor	Novice and expert users noted usability challenges due to weight, but confidence and workload improved
[68]	8	0	✓	✓	✓	A hallway in the Boai Campus BIO–ICT Building on the campus of National Yang Ming Chiao Tung University, Taiwan	Participants found the proposed route easy to navigate, with low perceived difficulty and medium confidence. Most intend to use it again.
[74]	20	0	✓	•	✓	Office area and simulated outdoor scenario	Positive feedback on usability and navigation; desire for detailed tutorials; satisfaction with daily use, challenges with multi-floor navigation.
[75]	4	0	✓	•	✓	In a corridor	All subjects found haptic instructions intuitive, enhancing safety and reducing hesitation compared to audio, though some suggested design improvements.
[77]	0	7	✓	•	✓	Various indoor places	Users find the wayfinding function useful, but discomfort due to weight is a significant concern.
[82]	3	0	✓	✓	•	Two different crosswalks	Intuitive, easy-to-understand verbal instructions, enhances street crossing safety
[87]	7	0	✓	✓	•	In unfamiliar building	High satisfaction with PathFinder's navigation assistance, intersection detection, and audio feedback.
[88]	0	6	✓	•	✓	In a configurable 12ft $\times$ 17ft room	High confidence, ease of use, and performance rated positively; verbal overview and haptics well-received for navigation assistance.

Future research should focus on the development of robust algorithms for continuous map updates and maintenance, including strategies for handling environmental changes over time. These strategies are essential to ensure that the navigation system can maintain its accuracy and reliability over extended periods, even in the face of environmental changes. By continuously updating and refining the map, the system can provide visually impaired users with up-to-date and relevant information about their surroundings, enabling them to navigate safely and confidently.

Deep learning integration The integration of deep learning with the SLAM algorithms for BVI navigation requires further investigation. Deep learning offers a versatile approach for enhancing various aspects of SLAM such as precise pose estimation under challenging conditions, relocalization, and loop-closure detection. This is important because deep learning can potentially improve the accuracy and robustness of SLAM in complex and dynamic environments, where traditional SLAM algorithms may struggle. For instance, deep learning can be used to improve feature detection and matching in low-light conditions or to predict and adapt to changes in the environment. Despite challenges, such as the need for large, accurately labeled datasets, the black-box nature of deep-learning models, and the computational intensity, the association between deep learning and SLAM holds promise for advancing navigation solutions for the visually impaired, particularly in challenging scenarios. The potential benefits of deep learning for SLAM are substantial, and overcoming these challenges could lead to significant advancements in assistive navigation technology.

Future research should focus on developing more efficient deep learning models that can operate effectively with limited computational resources and real-time constraints. This is crucial because visually impaired individuals need real-time feedback and guidance to navigate safely and effectively. Deep learning models that are computationally intensive or

TABLE 32. User satisfaction evaluation. This table includes studies that conducted user-based evaluations with actual participants to assess the effectiveness and satisfaction of the proposed systems - Part II.

Ref.	# Participants		Method of evaluation			Experimental site	User satisfaction
	BVI	Blindfolded	User study	Interview	Survey		
[91]	0	5	✓	•	•	In a laboratory	The significant improvement in success rate and task completion time, from 32% to 96% and 29.1s to 15.6s respectively, demonstrates the effectiveness of the proposed solution in aiding wayfinding and object manipulation.
[95]	14	0	✓	•	✓	A short route in a controlled and long route in a real-world public space.	Participants expressed high levels of contentment with the system's usability, effectiveness, and overall user experience in the study.
[97]	0	1	✓	•	•	Not mentioned	Assessed through successful task completion, yet occasional false positives slightly affect confidence.
[110]	10	1	By sighted	With BVI	•	In the Rhodes Research Center at Clemson University	An online interview with 10 BVI individuals via Zoom guided the choice of a 3D perception-enabled mobile platform with a speech-auditory interface.

require powerful hardware may not be practical for real-world use. Additionally, creating large-scale, accurately labeled datasets tailored for BVI navigation is crucial for training robust models. The lack of such datasets is a major obstacle to the development of effective deep-learning-based SLAM systems for visually impaired navigation.

Addressing the interpretability of deep learning models can also enhance the trust and transparency in these systems. Collaboration among machine learning experts, roboticists, and vision scientists can drive the development of innovative algorithms that leverage deep learning to enhance the reliability and accuracy of SLAM-based navigation aids for the visually impaired. This collaboration is essential to bring together the diverse expertise needed to develop effective and practical solutions.

Indoor and outdoor navigation integration Seamless transitions between indoor and outdoor environments are crucial for enhancing the independence and mobility of BVI individuals. However, most studies in our SLR have focused primarily on indoor environments. This limitation arises because indoor environments are often more structured and predictable than outdoor environments, making them easier to map and navigate using SLAM. Outdoor environments, on the other hand, present challenges such as varying lighting conditions, weather changes, and a wider range of obstacles.

Future research should aim to develop solutions that provide unified and consistent navigation experience in both indoor and outdoor settings. This is important because visually impaired individuals need to be able to navigate seamlessly between different environments in their daily lives. A navigation system that only works indoors or outdoors would be of limited use. Researchers should explore the integration of robust sensor fusion techniques and adaptive algorithms capable of handling the different conditions and challenges of these environments. Sensor fusion can combine data from multiple sensors, such as cameras, LiDAR, and IMUs, to provide a more comprehensive and accurate understanding of the environment. Adaptive algorithms can adjust the SLAM system's parameters in real time to account for changes in lighting, weather, and other environmental factors.

Obstacle detection Achieving detailed knowledge of obstacles and their characteristics is essential for BVI people. Although some studies included in the SLR addressed obstacle detection, the depth and accuracy of the obstacle information provided may still be limited. This limitation stems from the fact that traditional obstacle detection methods often focus on identifying the presence and location of obstacles but may not provide detailed information about their shape, size, or material, which is crucial for visually impaired individuals to make informed decisions during navigation.

To address the need for more detailed and accurate obstacle information, future research should focus on advancing SLAM algorithms to deliver context-aware obstacle detection. This involves integrating semantic understanding with precise spatial measurements, allowing the system to identify and interpret the nature and significance of obstacles accurately. By incorporating semantic understanding, SLAM systems can differentiate between different types of obstacles, such as curbs, stairs, or low-hanging branches, and provide more relevant and actionable information to the user. Precise spatial measurements are essential for accurately estimating the distance, size, and shape of obstacles, enabling visually impaired individuals to navigate safely around them.

In addition, it is crucial to develop algorithms that can learn and adapt to various obstacle types and scenarios. Drawing inspiration from the approaches used in robotics and autonomous drones, such as the real-time metric-semantic SLAM demonstrated by [118], can provide valuable insights. These approaches have demonstrated the feasibility and effectiveness of integrating semantic understanding with precise spatial measurements in real-time SLAM systems. Therefore, future research should prioritize improving both the depth and accuracy of obstacle information, while ensuring robust real-time performance and adaptability to various real-world conditions.

Semantic information integration Integrating semantic

TABLE 33. User satisfaction evaluation. This table includes studies that primarily conducted technical tests without direct user-based evaluations.

Ref.	# Participants		Method of evaluation			Experimental site	User satisfaction
	BVI	Blindfolded	User study	Interview	Survey		
[61]	The technical test was conducted in the corridor environment of a typical office building. Real-world user satisfaction was not directly assessed.						
[63]	78	0	•	•	✓	Surveyed BVI online using LymeSurvey to understand social networking needs	Not mentioned; the focus is on system development
[64]	0	1	✓	•	•	A laundry room, H-shaped hallway, classroom, and T-shaped hallway	Not mentioned; the focus is on system development
[65]	Not a user-based evaluation; the focus is on a technical test.						
[66]	Not explicitly mentioned; the focus is on a technical test.						
[69]	Not explicitly mentioned					Laboratory	Not mentioned
[70]	The focus is on a technical test					The Engineering East Hall of Virginia Commonwealth University	Technical focus only
[71]	Focuses on technical performance; real-world user satisfaction not directly assessed.						
[72]	The technical test was conducted in accommodation and office buildings. Real-world user satisfaction was not directly assessed.						
[73]	The technical test was conducted in a university lobby. Real-world user satisfaction was not directly assessed.						
[76]	Unspecified	0	✓	•	•	Not mentioned	Not mentioned
[78]	The technical test was conducted on seven datasets collected internally. Real-world user satisfaction was not directly assessed.						
[79]	The test was conducted in the College of Computer and Information Sciences at King Saud University, but user satisfaction was not directly assessed.						
[80]	1	5	Data was collected by participants to simulate indoor navigation, followed by offline analysis.				
[83]	The technical test was conducted in various indoor environments. Real-world user satisfaction was not directly assessed.						
[84]	The technical test was conducted in a corridor environment. Real-world user satisfaction was not directly assessed.						
[85]	The technical test was conducted in an office room and on the KITTI dataset. Real-world user satisfaction was not directly assessed.						
[89]	The technical test was conducted on the KITTI-02 dataset. Real-world user satisfaction was not directly assessed.						
[90]	The technical test included TUM RGB-D, Bonn RGB-D datasets, and real-life sequences, but didn't directly assess user satisfaction.						
[92]	3; Visual ability unspecified		✓	•	•	Not explicitly mentioned	User satisfaction not directly assessed; technical tests show high success rates and accuracy.
[93]	1; Visual ability unspecified		✓	•	•	In a laboratory	User satisfaction not assessed
[96]	Initiating target user testing was not feasible as the system was in the experimental stage.						
[98]	5	0	✓	•	•	Smith-Kettlewell building	Not mentioned
[99]	The technical test was conducted in the East Engineering Building at VCU. Real-world user satisfaction was not directly assessed.						
[100]	0	1	✓	•	•	In a laboratory	The technical test was conducted. Real-world user satisfaction was not directly assessed.
[101]	Paper develops a Reinforcement Learning environment to create a navigation assistant tailored for the BVI, without user-based evaluation.						
[102]	Technical tests conducted on the Karlsruhe dataset, indoor recorded dataset, and in a house; user satisfaction not assessed.						
[103]	Technical tests conducted in a research building; user satisfaction not assessed.						
[104]	Not explicitly mentioned; the focus is on a technical test.						
[105]	1; Visual ability unspecified		Not explicitly mentioned; the focus is on a technical test.				
[106]	The technical test was conducted on seven datasets collected in two buildings. Real-world user satisfaction was not directly assessed.						
[107]	0	Unspecified	✓	•	•	In various indoor environments (university campus, hotel, office building)	Not assessed
[108]	The technical test was conducted in an office and pedestrian street. Real-world user satisfaction was not directly assessed.						
[109]	The technical test was conducted in a laboratory. Real-world user satisfaction was not directly assessed.						
[111]	Real-world user satisfaction was not directly assessed.						
[112]	Technical tests conducted in a corridor; user satisfaction not assessed.						

information into SLAM algorithms can significantly enhance the performance and robustness of navigation systems for BVI individuals. This information can be used to refine the mapping and localization processes and enhance the overall reliability of navigation in complex environments. For instance, semantic information aids in rejecting outliers during loop-closure detection, which is a crucial SLAM step that identifies and matches previously visited locations. To advance this area, future research should focus on developing advanced techniques for semantic data extraction and integration within the SLAM frameworks. This is necessary because current methods for semantic data extraction and integration may not be efficient enough for real-time SLAM in complex environments. By developing more advanced techniques, researchers can improve the quality and reliability of semantic information used in SLAM, leading to better navigation performance.

Researchers should also explore methods to ensure real-time performance while maintaining the accuracy and detail of the semantic information. Real-time performance is crucial for providing visually impaired users with timely and relevant feedback during navigation. However, processing and integrating semantic information is computationally expensive. Therefore, it is important to develop methods that can balance real-time performance with the accuracy and detail of semantic information.

Realistic dataset creation Another crucial area for future research is the development of realistic and comprehensive datasets tailored for BVI navigation. Although some datasets exist for various SLAM applications, they often do not capture the unique challenges faced by the visually impaired, such as the need to navigate crowded spaces and avoid obstacles at different heights. Future research should focus on creating large-scale, diverse datasets that include various

TABLE 34. System prototype information for wearable devices - Part I.

Ref.	Functionalities	Sensors	Computing resource	HCI	Assistive tool	Battery life	ML-based
[57]	Navigation: positioning and wayfinding. Multi-target recognition: object localization, face recognition, and scene text recognition.	RGB-D camera, Ultrasonic sensor	Embedded computer, remote server	Audio, haptic	Electronic glasses and leg-mounted haptic modules	Over 12 hours under typical usage	Multi-target recognition
[58]	Real-time perception, remote assistance (leveraging WebRTC protocol), live broadcasts, chatrooms, real-time tagging	A depth, an RGB, and four gray scale cameras, an IMU	Microsoft Hololens2 device, GPU	Audio	Microsoft Hololens2		
[60]	Visual-based localization, location estimation, direction estimation, and navigation support	Phone's camera	Cloud server and Nvidia Jetson AGX Xavier	Audio	Android application		NetVLAD for global descriptors and SuperPoint for local descriptors to aid in the localization process
[62]	Active navigation with sub-goal inference, context-aware Object Relation Prior Knowledge	RealSense D435i RGB-D camera	Nvidia Jetson AGX Xavier	Audio	Helmet; white cane for obstacle avoidance		An unbiased Scene Graph Generation (SGG) model to create scene graphs, then aggregates them into an Object Relation Knowledge Graph
[63]	Scene description, face recognition, optical character recognition, obstacle recognition, social networking, remote assistance	IMU, stereo, IR (depth) camera	Raspberry Pi4, Cloud server	Audio, tactile	Smartphone		Faster RCNN: object detection, LSTM RNN: scene description, imitation-learning deep neural networks: navigation, Google Tensorflow im2txt: scene captioning, Neurotechnology's Verilook 12.2: face recognition
[66]	Real-time navigation, real-time semantic understanding, voice interaction, precise localization, real-time map generation	RGB-D camera	High-performance portable processor, cloud server	Audio	A forehead mounted camera and an earphone to obtain the output information		ENet for pixel-level semantic segmentation
[69]	Mapping, path planning, obstacle avoidance, transparent object detection, and path following.	RGB-D camdera (Asus Xtion Pro live), Ultrasonic sensor	Raspberry Pi3 B+	Audio	Head-mounted camera		•
[71]	Obstacle avoidance, scene perception, and hierarchical localization services.	RealSense RGB-D-IR camera, IMU, a customized GNSS receiver	A portable computer, Nvidia Jetson TX2	Not provided	Auxiliary glasses		NetVLAD and Dense Desc for advanced descriptor extraction.
[73]	Object detection and people detection for obstacle avoidance.	Two monocular cameras	Not mentioned	Audio	Smart eyeglasses		YOLOv2 for person detection.
[74]	Obstacle avoidance, surrounding perception.	RGB-D camera, IMU, GPS	A smartphone with Qualcomm Snapdragon 820 CPU 2.0 GHz	Audio	Eyeglasses-mounted sensors + smartphone		PeleeNet + SDD for object recognition
[75]	Adaptive artificial potential field path planning and semantic understanding.	Smartphone's camera, Gravity sensor, Ambient light sensor, Proximity sensor, Gyroscope Compass	A HUAWEI P20 smartphone with Kirin 970 CPU	Audio, haptic	ARCore-supported smartphone + haptic gloves		Not mentioned
[76]	Locating, way-finding, route following, and obstacle detection.	Fisheye and depth camera, ultrasonic rangefinder	An embedded CPU board	Audio, visual hints	Optical see-through glasses		•
[82]	Scene understanding, localization, object detection, path planning, path following, timely completion	Realseanse D435i RGB-D camera, compass sensor (BNO055 Bosch)	Nvidia Jetson Xavier AGX, NVidia	Audio	Head-worn camera		Bisenet & HarDNet for crosswalk, the end of the crosswalk (a red texture plate), and the crosswalk signal detection

TABLE 35. System prototype information for wearable devices - Part II.

Ref.	Functionalities	Sensors	Computing resource	HCI	Assistive tool	Battery life	ML-based
[85]	Detecting and locating objects of interest; guiding users efficiently to target objects.	Monocular camera	Nvidia Jetson Xavier NX Developer kit	Audio, Virtual Touch [122]	Chest-mounted camera		Pretrained YOLOv5 for object detection
[90]	Localization and mapping in dynamic environments, obstacle avoidance, dynamic object tracking.	RGB-D camera	Laptop	Audio	Smart glasses		PanopticFCN for obtaining the prior dynamic object information, OpenPose for obtaining a more accurate speed estimation of dynamic moving people
[91]	Locating a target object, wayfinding, motion guidance, and grasping the object.	Occipital-Structure Core sensor with a built-in Bosch BMI055 IMU, a color camera, and a global shutter stereo IR camera	Google Pixel 3 smartphone equipped with a Snapdragon 845 processor and 4 GB RAM.	Audio, haptic	Hand-worn device		TensorFlow Lite Object Detection API (MobileNet SSD model) for detecting the target object
[92]	Indoor navigation, real-time path planning, object of interest detection	RealSense D435i camera	Embeded Jetson nano 4GB, remote server based on Intel i7-8700 CPU, nvidia GTX1080 GPU, 64 GB DDR4 RAM	Audio, haptic	Smart E-glasses	Over 12 hours per charge	MobilenetV3-Yolov4-Lite, based on YOLOv4 and MobileNetV3 for object detection
[93]	Detecting and describing objects in environments, personalizing navigation through interactive dialogues and re-training, and locating users.	Google glasses camera	GPU server	Not mentioned	Google glasses		YOLOv4 and SSD for detecting and describe objects, a classical attention-based encoder-decoder model with LSTM and ResNet [123] for image captioning.
[97]	Scene perception, obstacle avoidance, and localization	RealSense R200 camera	Nvidia Jetson AGX Xavier processor	Audio	Smart glasses		RFNet for generating semantic labels
[100]	Semantic mapping and path planning, obstacle avoidance, environment perception	RPLIDAR A2, Microsoft Kinect VI, ZED stereo camera	Laptop	Audio	TurtleBot2 robot (to be replaced by portable device) + wearable sensors		YOLOv3 for landmark detection, Places365 for place recognition
[105]	Tracking blind pedestrians' paths	Hero3+ GoPro	Not mentioned	Not mentioned	Wearable camera		•
[108]	Obstacle avoidance, OCR, path planning, and human assistance via web application.	Stereo camera	Cloud server	Audio	Helmet-mounted camera + android-based smartphone		Recurrent Convolutional Neural for object detection and recognition, [124], [125] for scene parsing, [126], [127] for Optical Character Recognition, [128] for currency recognition, and [129] for traffic light recognition
[109]	Route guidance to a destination, obstacle avoidance	Monocular camera	A single CPU	Not mentioned	Wearable camera		•
[111]	Wayfinding and identifying short-term impediments with GeoNotify smartphone software.	Kinect camera	Not mentioned	Audio, haptic	Wearable camera		YOLOv4 Tiny for object detection

TABLE 36. System prototype information for handheld devices - Part I.

Ref.	Functionalities	Sensors	Computing resource	HCI	Assistive tool	Battery life	ML-based
[64]	Vibrations and sounds for obstacle avoidance, detailed mapping, real-time object recognition, and a smart cane for spatial orientation	RealSense RGB-D camera	Raspberry Pi	Audio, tactile	Smart cane		YOLO: Object detection
[67]	Obstacle avoidance, waypoint following, indoor/outdoor navigation, key object detection, user guidance through challenges	2D LiDAR, camera, GPS antenna, IMU	A portable microcontroller	vibrotactile, Audio, Grounded kinesthetic	Augmented cane	Microcontroller: 4.2 and motor: 5.2 hours	YOLOV3Tiny for object detection, a linear regression model for distance estimation
[70]	Active steering for user guidance, obstacle avoidance, and wayfinding.	Realsense D435 (RGB-D) Camera, VN100 IMU	UP Board computer	Audio, tactile	Smart cane		•
[72]	Accurate object detection, semantic mapping, and indoor localization services.	hand-held RGB-D camera	Odroid XU3 board, remote server (data processing)	Audio	No device		ConvNet for semantic information extraction and location inference; Inception-v3 for object recognition.
[77]	Wayfinding 3D object detection	SwissRanger SR4000 3D camera	Client: HP Stream 7 tablet, server: Lenovo ThinkPad T430 laptop (Intel i5-3320M 2.6GHz CPU, NVS 5400m with 96 CUDA core)	Audio	Smart cane		•
[78]	Pose estimation, obstacle detection and avoidance, wayfinding	SwissRanger SR4000 camera, IMU (VN-100 of VectorNav Technolofgies)	Up Board computer	Audio	Computer-vision-enhanced white cane		•
[80]	Wayfinding and localization.	iPhone 11 Pro sensors	iPhone 11 Pro	Not mentioned	iPhone 11 Pro		YOLOv2 for object detection to facilitate effective localization
[83]	Safely navigate to destinations in static unfamiliar areas and identify surrounding objects.	Neato XV-11 LiDAR, ultrasonic sensor, Raspberry pi camera (CameraPi)	Raspberry Pi3 B+	Audio	Sensors attached to a white cane	Long battery life expected	Tiny YOLOv2 for predicting object class.
[87]	Intersection detection and sign recognition	360° LiDAR, iPhone 12 Pro camera	Nvidia RTX 3080 graphic board	Audio, Handle interface	Suitcase-shaped robot	Low: 2.6 hours	EasyOCR and YOLOv5 for sign recognition
[88]	Finding socially preferred chairs	RGB-D from the RealSense D455 and IMU from the T265 cameras	Dell G15 laptop with an RTX 3060 GPU	Audio, haptic	Robotic cane	-	Detectron2 for object detection and Mask-RCNN for obtaining masks for classification.
[89]	Multi-sensory guidance, obstacle avoidance, real-time target detection	RBG-D camera	Uzel US-M5422 edge server, Raspberry Pi 4B	Audio, tactile	Smart cane		YOLO for target detection
[95]	Collision risk prediction, directional guidance, mode switching, obstacle avoidance, real-time feedback.	Two RealSense D435 RGB-D cameras, IMU, LiDAR	Laptop (Intel Core i7-8750H CPU @ 2.20GHz, NVIDIA GeForce GTX 1080 Mobile GPU)	Audio, tactile	A rolling suitcase-shaped device		YOLOv3 for detecting surrounding pedestrians.
[98]	Real-time localization and turn-by-turn directions.	iPhone 8's IMU and rear-facing camera	iPhone 8	Audio	Smartphone	Low: 13% usage in 16 minutes.	Not explicitly mentioned
[99]	Wayfinding, human intent detection, and human-robot interaction	RealSense D435 Camera, IMU (VN100 of VectorNav Technologies, LLC)	UP Board computer	Audio, motorized rolling tip	White-cane mounted camera		•

TABLE 37. System prototype information for handheld devices - Part II.

Ref.	Functionalities	Sensors	Computing resource	HCI	Assistive tool	Battery life	ML-based
[102]	Real-time localization, navigation, object detection, and distance-depth estimation, using a single monocular camera.	Monocular camera	Intel i7 processor	Audio	Smartphone		ACF detector for object detection to identify trained objects of interest for localization.
[106]	Pose estimation, wayfinding assistance	Time-of-flight camera (SwissRanger SR4000), IMU (VN 100 of VectorNav Technologies, LLC)	UP Board computer	Not mentioned	Computer-vision-enhanced white cane		•
[107]	Indoor mapping, path planning, control panel interface, and object avoidance.	Wide-angle lens camera, gyroscope, accelerometers, and infrared sensor on Google Tango, IMU on the cane.	Google Tango, microcontroller	Audio, haptic	Smart cane + Google Tango	Lasts approximately 17 hours	•
[110]	Global path finding, local path re-planning, and obstacle avoidance.	Camera, IMU, inbuilt 3D LiDAR from iPhone 12 Pro Max	Smartphone, AWS Lambda on cloud	Audio	iPhone 12 Pro Max		SFSpeechRecognizer from iOS for speech-to-text, ResNet for extracting feature representation of each viewpoint for scene-graph map construction, EnvDrop for path exploration, and reinforcement learning for training the Vision-Language Navigation agent to navigate indoor environments based on language instructions.

TABLE 38. Prototype information for Robot systems, ride-on systems, and setups without specific devices or where device information is not mentioned.

Ref.	Functionalities	Sensors	Computing resource	HCI	Assistive tool	Battery life	ML-based
				Robot systems			
[59]	Considerate navigation, spatial risk mapping, adaptive motion control	A 2D range sensor, 2 RGB-D RealSense D435 cameras	Two notebook PCs	Audio	Wheeled guide mobile robot		Pedestrian detection (OpenPose), OpenCV ObjDetect Module Face Recognition, Yolact obstacle recognition, Speech recognition
[61]	Localization	LiDAR, cameras	Nvidia Titan X GPU	No device	No device		GAN-based localization
[68]	UWB beacons for audio-based environmental information, dynamic obstacle avoidance, wall following, adjustable speed, and emergency stop.	Velodyne LiDAR VLP16, RealSense D345 depth camera	an Intel NUC computer, Nvidia Jetson TX2, Raspberry Pi3	Audio, haptic	Guiding robot		Reinforcement learning
[79]	Robot navigation, obstacle avoidance, path planning, and user interaction	Encoder, IMU, laser distance sensor, camera	Raspberry Pi 3 Model B and B+	Audio	Smart robot (Turtlebot3 robot)		•
				Ride-on systems			
[104]	Autonomous navigation, real-time mapping and localization, obstacle avoidance, accurate steering.	IMU, MPU-9250, stereo camera, laser, LiDAR	Nvidia Jetson TX2, Raspberry Pi, and Arduino.	Steering control	Intelligent autonomous scooter		•
[112]	Navigation, path following, obstacle avoidance.	Hokuyo's URG-04LX-UG01 Laser Range Finder (LRF) sensor, Microsoft LifeCam HD-5000 USB camera, MTi-30 Attitude Heading Reference System (AHRS) IMU sensor from Xsens	PC	Autonomous navigation	Person carrier robot (wheelchair)		•
				Unspecified setups			
[56]	Navigation	Camera	Smartphone	Audio	Smartphone		fill this cell
[65]	No prototype implemented; the proposed approach for trajectory forecasting was tested with a robot.						AlphaPose: to detect and track the nearby people appearing in each frame; PSPNet: to segment scene semantics; Monodepth2: to estimate depth from the monocular RGB frames; Transformer-based encoder-decoder neural network model: containing a novel cascaded cross-attention mechanism to fuse encodings of different modalities for trajectory forecasting.
[84]	Real-time global localization	ZED2 RGB-D, IMU	Nvidia Jetson AGX Xavier microprocessor	No device	No device		MobileNetV2 with PPM for constructing semantic point cloud
[96]	Route planning and obstacle avoidance.	RGB-D camera, IMU, LiDAR	Not mentioned	Audio	No device		Not mentioned
[101]	Paper develops a Reinforcement Learning environment to create a navigation assistant tailored for the BVI community, without prototype implementation.						
[103]	Enhanced indoor localization and navigation using environmental texts.	Smartphone and ZED cameras	Nvidia Jetson TX2	No device	No device	No device	CNN for text detection and recognition

TABLE 39. Classification of references based on functionalities

Functionality	References
Localization	[60], [61], [66], [72], [76], [78], [80], [82], [84], [90], [93], [97], [98], [102]–[104], [106]
Mapping	[64], [66], [69], [90], [104], [107]
Obstacle avoidance	[64], [67]–[71], [73], [74], [76], [78], [79], [89], [90], [95]–[97], [100], [104], [107]–[110], [112]
Object detection	[63], [64], [72], [73], [77], [82], [83], [87], [89], [93], [102]
Object localization	[57], [67], [71], [85], [88], [91], [92]
Scene understanding	[58], [63], [66], [71], [74], [75], [82], [97], [100]
Path planning	[69], [75], [79], [82], [92], [96], [100], [107], [108], [110]
Face recognition	[57], [63]
Remote assistance	[58], [63]
Social networking	[63], [108]
Way-finding	[57], [67], [70], [76]–[78], [80], [91], [99], [106], [109], [111]
Dynamic object tracking	[90]
Semantic mapping	[72], [100]
Collision prediction	[95]
Human intent detection	[99]
Spatial risk mapping	[59]

TABLE 40. Classification of papers based on HCI

HCI Mechanism	References
Audio	[56]–[60], [62]–[64], [66]–[70], [72]–[79], [82], [83], [85], [87]–[92], [95]–[100], [102], [107], [108], [110], [111]
Tactile	[63], [64], [67], [70], [89], [95]
Haptic	[57], [68], [75], [88], [91], [92], [107], [111]
Visual hints	[76]
Motorized rolling tip	[99]
Handle interface	[87]
Steering control	[104]
Virtual touch	[85]
Grounded Kinesthetic	[67]
Not Provided/Not mentioned	[61], [71], [80], [84], [93], [101], [103], [105], [106], [109]

indoor and outdoor settings, different lighting conditions, and dynamic elements. This is important because the lack of such datasets hinders the development and evaluation of SLAM algorithms that are specifically designed for BVI navigation. By creating datasets that reflect real-world challenges faced by visually impaired individuals, researchers can develop more effective and reliable navigation solutions.

Computing resources and battery life The development and deployment of SLAM-based navigation systems for the visually impaired face significant challenges related to computing resources and battery life. SLAM algorithms, particularly when integrated with deep learning models, often demand substantial computing power that can quickly drain battery life and generate heat in mobile devices. This is a critical issue because visually impaired individuals need portable and comfortable navigation aids that can operate for extended periods without overheating or frequent recharging.

Additionally, intensive computations and continuous sensor usage drain the battery life quickly, limiting the usability of the system in real-world daily scenarios. This limitation can significantly hinder the adoption and effectiveness of the SLAM-based navigation systems. Future research should focus on optimizing SLAM algorithms and deep learning models for low-power devices without compromising the accuracy or real-time performance. This is essential for developing energy-efficient solutions that can operate on portable devices with limited battery capacity, ensuring that visually impaired individuals can rely on these systems for extended periods without interruption. Exploring edge computing solutions, developing more efficient neural network architectures, and enhancing battery management techniques could help to address these challenges. These approaches can collectively contribute to reducing the computational burden and power consumption of SLAM-based navigation systems, making them more practical and sustainable for real-world use.

TABLE 41. Classification of papers based on assistive tool

Assistive Tool	References
Smartphone-based	[56], [60], [63], [74], [75], [80], [98], [102], [108], [110]
Glass-based	[57], [71], [73], [74], [76], [90], [92], [93], [97]
Cane-based	[64], [67], [70], [77], [78], [83], [88], [89], [99], [106], [107]
Robotic systems	[59], [68], [79], [87], [95]
Ride-on systems	[104], [112]
Wearable sensor	[85], [100], [105], [109], [111]
Head-worn sensor	[58], [62], [66], [82], [91], [105]
No specific device/ Unspecified	[61], [72], [84], [96], [101], [103]

Human-computer interaction for SLAM-based assistive devices An important area for future research is to improve the human-computer interaction (HCI) aspects of SLAM-based assistive devices for BVI individuals. Although SLAM techniques have shown great potential in gathering and processing environmental information, effectively communicating this information to BVI users remains a significant challenge. Future research should focus on developing intuitive and non-intrusive methods to convey complex spatial data and navigational instructions to BVI users. This includes:

- Multi-modal feedback systems: Exploring combinations of audio, haptic, and other non-visual feedback methods to provide rich contextual information without overwhelming the user.
- Adaptive interfaces: Developing interfaces that can adjust the level and type of information provided based on the user's preferences, familiarity with the environment, and the current situation.
- Natural language processing: Improving the ability of systems to understand and respond to natural language queries, allowing for a more intuitive interaction between the user and device.
- Cognitive load optimization: Investigating ways to balance the provision of detailed environmental information, ensuring that users receive the necessary guidance without cognitive overload.
- Real-time situational awareness: Developing methods to effectively communicate dynamic elements of the environment, such as moving obstacles or changing traffic conditions in real-time.

Addressing these HCI challenges will be crucial in translating the technical capabilities of SLAM into practical, user-friendly assistive devices that can significantly enhance the mobility and independence of BVI individuals. Future re-

TABLE 42. Categorization of machine learning techniques used in assistive solutions

Technique	References
Object detection	
YOLO	[64], [89]
YOLOv2	[73], [80], [83]
YOLOv3 (Tiny)	[67], [95], [100]
YOLOv4 (Tiny)	[93], [111]
YOLOv5	[85], [87]
Yolact obstacle detection	[59]
(Faster) RCNN	[63], [108]
Bisenet & HarDNet: crosswalk and signal detection	[82]
Detectron2	[88]
TensorFlow Lite API	[91]
MobileNetV3-Yolov4-Lite	[92]
ACF detector	[102]
CNN for text detection	[103]
RCNN	[108]
Object recognition	
Multi-target recognition	[57]
Inception-v3	[72]
PeleeNet + SDD	[74]
EasyOCR: sign recognition	[87]
Face recognition	
OpenCV ObjDetect Module Face Recognition	[59]
Neurotechnology's Verilook 12.2: face recognition	[63]
Semantic segmentation and scene understanding	
LSTM RNN: scene description	[63]
PSPNet: segment scene semantics	[65]
ENet: pixel-level semantic segmentation	[66]
MobileNetV2: constructing semantic point cloud	[84]
Mask RCNN	[88]
PanopticFCN	[90]
RFNet: generating semantic labels	[97]
Scene parsing using [124], [125]	[108]
ResNet: scene-graph map construction	[110]
Image captioning	
Google Tensorflow im2txt	[63]
Classical model with LSTM and ResNet [123]	[93]
Visual odometry and localization	
NetVLAD: global descriptors	[60], [71]
SuperPoint: local descriptors	[60]
GAN-based localization	[61]
Monodepth2: estimating depth from RGB frame	[65]
Transformer-based model: trajectory forecasting	[65]
Deep Descriptors	[71]
ConvNet: location inference	[72]
Places365: place recognition	[100]
Reinforcement learning & other techniques	
Reinforcement learning	[68], [101], [110]
Speech recognition	[59]
OpenPose: Pedestrian detection	[59], [90]
Scene Graph Generation	[62]
LSTM RNN	[63]
Imitation-learning DNN	[63]
AlphaPose: detecting and tracking pedestrians	[65]
Linear regression model: distance estimation	[67]
Optical Character Recognition	[108]
Currency recognition	[108]
SFSpeechRecognizer from iOS	[110]
EnvDrop: path exploration	[110]

search in this area should involve close collaboration with BVI users to ensure that the developed interfaces meet their needs and preferences.

Product development and collaboration Notably, all reviewed approaches were prototypes in the early stages of research and are not yet practical. This might be due to the absence of a unified community or group dedicated to solving the BVI navigation challenges. Much of the work in this domain has been conducted by academic groups or small companies that often fail to produce feasible final products. This underscores a significant future opportunity to develop collaboration and to bridge the gap between research and practical implementation.

Additionally, efforts should be made to develop standardized evaluation metrics and protocols to ensure that the developed systems meet real-world needs and can be effectively transitioned from prototypes to market-ready solutions. Standardized evaluation metrics and protocols are essential to ensure that assistive navigation systems are evaluated consistently and objectively. This can help identify the strengths and weaknesses of different approaches and guide the development of more effective solutions. Encouraging partnerships with technology companies can also accelerate the commercialization process. These partnerships provide the necessary support to bring innovative solutions to the market.

In conclusion, the future of SLAM for the visually impaired navigation is promising. Continued research efforts have the potential to develop SLAM algorithms tailored for BVI navigation, empowering visually impaired individuals with a safe and independent means of navigating their surroundings.

V. CONCLUSION

This study presents a systematic literature review of recent studies on SLAM-based solutions for BVI navigation. Excluding papers published before 2017, this review focused on the latest advancements, innovations, and considerations, resulting in a more relevant and comprehensive understanding of the current state of research. The insights provided by this systematic literature review are intended to guide researchers in the academic and research communities. They inform the existing gaps and future opportunities to address the challenges faced by SLAM-based assistive solutions.

Relevant data were extracted from 54 selected studies that adhered to the SLR selection criteria to address the research questions. By analyzing the selected papers based on their SLAM techniques, we observed that the majority of the studies utilized visual SLAM techniques, such as ORB-SLAM3, owing to their advantages for visual sensors.

Several studies have introduced novel strategies for addressing localization and mapping challenges tailored to the specific requirements of their research, whereas certain studies have employed existing spatial tracking frameworks to develop navigation solutions. We also investigated the advantages and limitations of the SLAM techniques, as highlighted in the studies under review. Notably, most studies have leveraged accurate localization features of SLAM.

We investigated the challenging scenarios encountered by SLAM-based navigation systems, which have been addressed in the literature. Additionally, we discussed practical challenges and considerations that affect the usability and adoption of these systems. Furthermore, we analyzed how the proposed SLAM-based solutions improve the mobility and navigation of visually impaired individuals. We evaluated the effectiveness of these solutions in real-world scenarios and assessed the user satisfaction to understand their practical impact on BVI mobility. Finally, we identified gaps, opportunities, and areas of interest that could be explored further in future research, such as addressing challenges in crowded environments, improving real-world applicability, integrating deep learning, and ensuring long-term navigation effectiveness in SLAM-based solutions for visually impaired navigation.

Given the widespread application of SLAM in robotic, autonomous drones, and auto-driving car navigation, these techniques can be adapted to ensure safe and independent BVI navigation. This is particularly important in dynamic and challenging environments, including those with varying lighting conditions where research opportunities remain abundant. The potential of integrating these techniques into the navigation of visually impaired individuals continues to be an open and promising avenue.

ACKNOWLEDGMENT

We would like to extend our sincere gratitude to Giovanni Cioffi, whose insightful feedback and thorough review significantly contributed to the refinement of this manuscript.

REFERENCES

[1] P. K. Panigrahi and S. K. Bisoy, "Localization strategies for autonomous mobile robots: A review," *Journal of King Saud University-Computer and Information Sciences*, 2021.

[2] Y. Alkendi, L. Seneviratne, and Y. Zweiri, "State of the art in vision-based localization techniques for autonomous navigation systems," IEEE Access, vol. 9, pp. 76847-76874, 2021.

[3] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309–1332, 2016.

[4] M. S. A. Khan, D. Hussain, K. Naveed, U. S. Khan, I. Q. Mundial, and A. B. Aqeel, "Investigation of Widely Used SLAM Sensors Using Analytical Hierarchy Process," Journal of Sensors, vol. 2022, 2022.

[5] S. Real and A. Araujo, "Navigation systems for the blind and visually impaired: Past work, challenges, and open problems," Sensors, vol. 19, no. 15, p. 3404, 2019.

[6] M. D. Messaoudi, B.-A. J. Menelas, and H. Mcheick, "Review of Navigation Assistive Tools and Technologies for the Visually Impaired," Sensors, vol. 22, no. 20, p. 7888, 2022.

[7] A. T. Parker, M. Swobodzinski, J. D. Wright, K. Hansen, B. Morton, and E. Schaller, "Wayfinding tools for people with visual impairments in real-world settings: a literature review of recent studies," in Frontiers in Education, vol. 6, p. 723816, 2021.

[8] S. Khan, S. Nazir, and H. U. Khan, "Analysis of navigation assistants for blind and visually impaired people: A systematic review," IEEE Access, vol. 9, pp. 26712-26734, 2021.

[9] M. R. M. Romlay, S. F. Toha, A. M. Ibrahim, and I. Venkat, "Methodologies and evaluation of electronic travel aids for the visually impaired people: a review," Bulletin of Electrical Engineering and Informatics, vol. 10, no. 3, pp. 1747-1758, 2021.

[10] X. Zhang, X. Yao, L. Hui, F. Song, and F. Hu, "A Bibliometric Narrative Review on Modern Navigation Aids for People with Visual Impairment," Sustainability, vol. 13, no. 16, p. 8795, 2021.

[11] K. Manjari, M. Verma, and G. Singal, "A survey on assistive technology for visually impaired," Internet of Things, vol. 11, p. 100188, 2020.

[12] Md M. Islam, M. S. Sadi, K. Z. Zamli, and M. M. Ahmed, "Developing walking assistants for visually impaired people: A review," IEEE Sensors Journal, vol. 19, no. 8, pp. 2814-2828, 2019.

[13] H. Fernandes, P. Costa, V. Filipe, H. Paredes, and J. Barroso, "A review of assistive spatial orientation and navigation technologies for the visually impaired," Universal Access in the Information Society, vol. 18, no. 1, pp. 155-168, 2019.

[14] M. D. Messaoudi, B.-A. J. Menelas, and H. Mcheick, "Review of Navigation Assistive Tools and Technologies for the Visually Impaired," Sensors, vol. 22, no. 20, pp. 7888, 2022.

[15] D. Khan, Z. Cheng, H. Uchiyama, S. Ali, M. Asshad, and K. Kiyokawa, "Recent advances in vision-based indoor navigation: A systematic literature review," Computers & Graphics, 2022.

[16] J. Wang, E. Liu, Y. Geng, X. Qu, and R. Wang, "A Survey of 17 Indoor Travel Assistance Systems for Blind and Visually Impaired People," IEEE Transactions on Human-Machine Systems, vol. 52, no. 1, pp. 134-148, 2021.

[17] D. Plikynas, A. Žvironas, M. Gudauskis, A. Budrionis, P. Daniušis, and I. Sliesoraitytė, "Research advances of indoor navigation for blind people: A brief review of technological instrumentation," IEEE Instrumentation & Measurement Magazine, vol. 23, no. 4, pp. 22-32, 2020.

[18] A. R. Façanha, T. Darin, W. Viana, and J. Sánchez, "O&M indoor virtual environments for people who are blind: A systematic literature review," ACM Transactions on Accessible Computing (TACCESS), vol. 13, no. 2, pp. 1-42, 2020.

[19] W. C. S. Simões, G. S. Machado, A. Sales, M. M. de Lucena, N. Jazdi, and V. F. de Lucena, "A review of technologies and techniques for indoor navigation systems for the visually impaired," Sensors, vol. 20, no. 14, p. 3935, 2020.

[20] D. Plikynas, A. Žvironas, A. Budrionis, and M. Gudauskis, "Indoor navigation systems for visually impaired persons: Mapping the features of existing technologies to user needs," Sensors, vol. 20, no. 3, p. 636, 2020.

[21] R. N. Kandalan and K. Namuduri, "Techniques for constructing indoor navigation systems for the visually impaired: A review," IEEE Transactions on Human-Machine Systems, vol. 50, no. 6, pp. 492-506, 2020.

[22] A. Zvironas, M. Gudauskis, and D. Plikynas, "Indoor electronic traveling aids for visually impaired: systemic review," in 2019 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 936-942, IEEE, 2019.

[23] H. Walle, C. De Runz, B. Serres, and G. Venturini, "A Survey on Recent Advances in AI and Vision-Based Methods for Helping and Guiding Visually Impaired People," Applied Sciences, vol. 12, no. 5, p. 2308, 2022.

[24] M. M. Valipoor and A. de Antonio, "Recent trends in computer vision-driven scene understanding for VI/blind users: a systematic mapping," Universal Access in the Information Society, pp. 1-23, 2022.

[25] Z. Fei, E. Yang, H. Hu, and H. Zhou, "Review of machine vision-based electronic travel aids," in 2017 23rd International Conference on Automation and Computing (ICAC), pp. 1-7, IEEE, 2017.

[26] S. Sivan and G. Darsan, "Computer vision-based assistive technology for blind and visually impaired people," in Proceedings of the 7th International Conference on Computing Communication and Networking Technologies, pp. 1-8, 2016.

[27] R. Jafri, S. A. Ali, H. R. Arabnia, and S. Fatima, "Computer vision-based object recognition for the visually impaired in an indoors environment: a survey," The Visual Computer, vol. 30, no. 11, pp. 1197-1222, 2014.

[28] P. Xu, G. A. Kennedy, F.-Y. Zhao, W.-J. Zhang, and R. Van Schyndel, "Wearable obstacle avoidance electronic travel aids for blind and visually impaired individuals: A systematic review," IEEE Access, 2023.

[29] M. Hersh, "Wearable travel aids for blind and partially sighted people: A review with a focus on design issues," Sensors, vol. 22, no. 14, p. 5454, 2022.

[30] A. D. P. Dos Santos, A. H. G. Suzuki, F. O. Medola, and A. Vaezipour, "A systematic review of wearable devices for orientation and mobility of adults with visual impairment and blindness," IEEE Access, vol. 9, pp. 162306-162324, 2021.

[31] R. Tapu, B. Mocanu, and T. Zaharia, "Wearable assistive devices for visually impaired: A state of the art survey," Pattern Recognition Letters, vol. 137, pp. 37-52, 2020.

[32] A. Chaudhary, Dr Verma, and others, "State of Art on Wearable Device to Assist Visually Impaired Person Navigation in Outdoor Environment," in Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE), 2019.

[33] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for the blind: a survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, no. 1, pp. 25-35, 2009.

[34] K. Thiyagarajan, S. Kodagoda, M. Luu, T. Duggan-Harper, D. Ritchie, K. Prentice, and J. Martin, "Intelligent Guide Robots for People who are Blind

or have Low Vision: A Review," Vision Rehabilitation International, vol. 13, no. 1, pp. 1-15, 2022.
[35] Md N. Alam, Md M. Islam, Md A. Habib, and M. B. Mredul, "Staircase detection systems for the visually impaired people: a review," International Journal of Computer Science and Information Security (IJCSIS), vol. 16, no. 12, pp. 13-18, 2018.
[36] A. Kinra, W. Walia, and S. Sharanya, "A Comprehensive and Systematic Review of Deep Learning Based Object Recognition Techniques for the Visually Impaired," in 2023 2nd International Conference on Computational Systems and Communication (ICCSC), pp. 1-6, IEEE, 2023.
[37] K. M. Reyes Leiva, M. Jaén-Vargas, B. Codina, and J. J. Serrano Olmedo, "Inertial measurement unit sensors in assistive technologies for visually impaired people, a review," Sensors, vol. 21, no. 14, p. 4767, 2021.
[38] F. E.-Z. El-Taher, A. Taha, J. Courtney, and S. Mckeever, "A systematic review of urban navigation systems for visually impaired people," Sensors, vol. 21, no. 9, p. 3103, 2021.
[39] G. Motta et al., "Overview of smart white canes: connected smart cane from front end to back end," in Mobility of Visually Impaired People, pp. 469-535, Springer, 2018.
[40] H. L. Tan, T. Aplin, T. McAuliffe, and H. Gullo, "An exploration of smartphone use by, and support for people with vision impairment: a scoping review," Disability and Rehabilitation: Assistive Technology, pp. 1-26, 2022.
[41] A. Budrionis, D. Plikynas, P. Daniušis, and A. Indrulionis, "Smartphone-based computer vision travelling aids for blind and visually impaired individuals: A systematic review," Assistive Technology, vol. 34, no. 2, pp. 178-194, 2022.
[42] A. Khan and S. Khusro, "An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities," Universal Access in the Information Society, vol. 20, no. 2, pp. 265-298, 2021.
[43] M. J. Grant and A. Booth, "A typology of reviews: an analysis of 14 review types and associated methodologies," Health Information & Libraries Journal, vol. 26, no. 2, pp. 91-108, 2009.
[44] D. Pati and L. N. Lorusso, "How to write a systematic review of the literature," HERD: Health Environments Research & Design Journal, vol. 11, no. 1, pp. 15-30, 2018.
[45] Staffs Keele and others, "Guidelines for performing systematic literature reviews in software engineering," Technical report, ver. 2.3 ebse technical report, ebse, 2007.
[46] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering–a systematic literature review," Information and Software Technology, vol. 51, no. 1, pp. 7-15, 2009.
[47] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," Information and Software Technology, vol. 64, pp. 1-18, 2015.
[48] A. Carrera-Rivera, W. Ochoa-Agurto, F. Larrinaga, and G. Lasa, "How-to conduct a systematic literature review: A quick guide for computer science research," MethodsX, p. 101895, 2022.
[49] C. Okoli and K. Schabram, "A guide to conducting a systematic literature review of information systems research," 2010.
[50] Y. Xiao and M. Watson, "Guidance on conducting a systematic literature review," Journal of Planning Education and Research, vol. 39, no. 1, pp. 93-112, 2019.
[51] A. Fink, Conducting Research Literature Reviews: From the Internet to Paper, Sage publications, 2019.
[52] B. Kitchenham et al., "Guidelines for performing systematic literature reviews in software engineering version 2.3," Engineering, vol. 45, no. 4ve, pp. 1051, 2007.
[53] D. Papaioannou, A. Sutton, and A. Booth, "Systematic approaches to a successful literature review," Systematic approaches to a successful literature review, pp. 1-336, 2016, SAGE Publications Ltd.
[54] H. Zhang, M. A. Babar, and P. Tell, "Identifying relevant studies in software engineering," Information and Software Technology, vol. 53, no. 6, pp. 625-637, 2011.
[55] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, pp. 1-10, 2014.
[56] D. Ahmetovic, F. Avanzini, A. Baratè, C. Bernareggi, M. Ciardullo, G. Galimberti, L. A. Ludovico, S. Mascetti, and G. Presti, "Sonification of navigation instructions for people with visual impairment," International Journal of Human-Computer Studies, vol. 177, p. 103057, 2023.
[57] G. Li, J. Xu, Z. Li, C. Chen, and Z. Kan, "Sensing and Navigation of Wearable Assistance Cognitive Systems for the Visually Impaired," IEEE Transactions on Cognitive and Developmental Systems, 2022.
[58] J. Song, J. Wang, S. Zhu, H. Hu, M. Zhai, J. Xie, and H. Gao, "Mixture reality-based assistive system for visually impaired people," Displays, vol. 78, p. 102449, 2023.
[59] B. Zhang, M. Okutsu, R. Ochiai, M. Tayama, and H. Lim, "Research on Design and Motion Control of a Considerate Guide Mobile Robot for Visually Impaired People," IEEE Access, 2023.
[60] A. Yang, M. Beheshti, T. E. Hudson, R. Vedanthan, W. Riewpaiboon, P. Mongkolwat, C. Feng, and J.-R. Rizzo, "UNav: An Infrastructure-Independent Vision-Based Navigation System for People with Blindness and Low Vision," Sensors, vol. 22, no. 22, p. 8894, 2022.
[61] G. Zhou, S. Xu, S. Zhang, Y. Wang, and C. Xiang, "Multi-Floor Indoor Localization Based on Multi-Modal Sensors," Sensors, vol. 22, no. 11, p. 4162, 2022.
[62] X. Hou, H. Zhao, C. Wang, and H. Liu, "Knowledge-driven indoor object-goal navigation aid for visually impaired people," Cognitive Computation and Systems, vol. 4, no. 4, pp. 329-339, 2022.
[63] D. Plikynas, A. Indriulionis, A. Laukaitis, and L. Sakalauskas, "Indoor-guided navigation for people who are blind: Crowdsourcing for route mapping and assistance," Applied Sciences, vol. 12, no. 1, p. 523, 2022.
[64] Z. Xie, Z. Li, Y. Zhang, J. Zhang, F. Liu, and W. Chen, "A multi-sensory guidance system for the visually impaired using YOLO and ORB-SLAM," Information, vol. 13, no. 7, p. 343, 2022.
[65] J. Qiu, L. Chen, X. Gu, F. P-W Lo, Y.-Y. Tsai, J. Sun, J. Liu, and B. Lo, "Egocentric human trajectory forecasting with a wearable camera and multi-modal fusion," IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 8799-8806, 2022.
[66] Z. Chen, X. Liu, M. Kojima, Q. Huang, and T. Arai, "A wearable navigation device for visually impaired people based on the real-time semantic visual SLAM system," Sensors, vol. 21, no. 4, p. 1536, 2021.
[67] P. Slade, A. Tambe, and M. J. Kochenderfer, "Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision," Science Robotics, vol. 6, no. 59, p. eabg6594, 2021.
[68] C.-L. Lu, Z.-Y. Liu, J.-T. Huang, C.-I. Huang, B.-H. Wang, Y. Chen, N. Wu, H.-C. Wang, L. Giarré, and P.-Y. Kuo, "Assistive navigation using deep reinforcement learning guiding robot with UWB/voice beacons and semantic feedbacks for blind and visually impaired people," Frontiers in Robotics and AI, p. 176, 2021.
[69] H. Hakim and A. Fadhil, "Indoor Wearable Navigation System Using 2D SLAM Based on RGB-D Camera for Visually Impaired People," in Proceedings of First International Conference on Mathematical Modeling and Computational Science, pp. 661-672, 2021.
[70] H. Zhang, L. Jin, and C. Ye, "An RGB-D Camera Based Visual Positioning System for Assistive Navigation by a Robotic Navigation Aid," IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 8, pp. 1389-1400, 2021.
[71] R. Cheng, W. Hu, H. Chen, Y. Fang, K. Wang, Z. Xu, and J. Bai, "Hierarchical visual localization for visually impaired people using multimodal images," Expert Systems with Applications, vol. 165, p. 113743, 2021.
[72] Q. Liu, R. Li, H. Huosheng, and D. Gu, "Indoor topological localization based on a novel deep learning technique," Cognitive Computation, vol. 12, pp. 528-541, 2020.
[73] S. Jin, M. U. Ahmed, J. W. Kim, Y. H. Kim, and P. K. Rhee, "Combining obstacle avoidance and visual simultaneous localization and mapping for indoor navigation," Symmetry, vol. 12, no. 1, p. 119, 2020.
[74] J. Bai, Z. Liu, Y. Lin, Y. Li, S. Lian, and D. Liu, "Wearable travel aid for environment perception and navigation of visually impaired people," Electronics, vol. 8, no. 6, p. 697, 2019.
[75] X. Zhang, X. Yao, Y. Zhu, and F. Hu, "An ARCore based user centric assistive navigation system for visually impaired people," Applied Sciences, vol. 9, no. 5, p. 989, 2019.
[76] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Virtual-blind-road following-based wearable navigation device for blind people," IEEE Transactions on Consumer Electronics, vol. 64, no. 1, pp. 136-143, 2018.
[77] H. Zhang and C. Ye, "An indoor wayfinding system based on geometric features aided graph SLAM for the visually impaired," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 9, pp. 1592-1604, 2017.
[78] H. Zhang and C. Ye, "Plane-aided visual-inertial odometry for 6-DOF pose estimation of a robotic navigation aid," IEEE Access, vol. 8, pp. 90042-90051, 2020.

[79] F. Albogamy, T. Alotaibi, G. Alhawdan, and F. Mohammed, "SRAVIP: Smart Robot Assistant for Visually Impaired Persons," International Journal of Advanced Computer Science and Applications, vol. 12, no. 7, 2021.
[80] R. Crabb, S. A. Cheraghi, and J. M. Coughlan, "A Lightweight Approach to Localization for Blind and Visually Impaired Travelers," Sensors, vol. 23, no. 5, p. 2701, 2023.
[81] D. S. Salih and A. M. Ali, "Appearance-based indoor place recognition for localization of the visually impaired person," ZANCO Journal of Pure and Applied Sciences, vol. 31, no. 4, p. 2412, 2019.
[82] H. Son and J. Weiland, "Wearable System to Guide Crosswalk Navigation for People With Visual Impairment," Frontiers in Electronics, vol. 2, article 790081, 2022.
[83] H. Hakim and A. Fadhil, "Indoor Low Cost Assistive Device using 2D SLAM Based on LiDAR for Visually Impaired People," Iraqi Journal for Electrical & Electronic Engineering, vol. 15, no. 2, 2019.
[84] RG Goswami, PV Amith, J Hari, A Dhaygude, P Krishnamurthy, J Rizzo, A Tzes, and F Khorrami, "Efficient Real-Time Localization in Prior Indoor Maps Using Semantic SLAM," in 2023 9th International Conference on Automation, Robotics and Applications (ICARA), pp. 299-303, 2023.
[85] Y. Hao, J. Feng, J.-R. Rizzo, Y. Wang, and Y. Fang, "Detect and Approach: Close-Range Navigation Support for People with Blindness and Low Vision," in European Conference on Computer Vision, pp. 607-622, 2022.
[86] S. Kayukawa, K. Higuchi, S. Morishima, and K. Sakurada, "3DMovieMap: an Interactive Route Viewer for Multi-Level Buildings," in Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1-11, 2023.
[87] M. Kuribayashi, T. Ishihara, D. Sato, J. Vongkulbhisal, K. Ram, S. Kayukawa, H. Takagi, S. Morishima, and C. Asakawa, "PathFinder: Designing a Map-less Navigation System for Blind People in Unfamiliar Buildings," in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1-16, 2023.
[88] S. Agrawal, M. E. West, and B. Hayes, "A novel perceptive robotic cane with haptic navigation for enabling vision-independent participation in the social dynamics of seat choice," in 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 9156-9163, 2022.
[89] C. Rui, Y. Liu, J. Shen, Z. Li, and Z. Xie, "A multi-sensory blind guidance system based on YOLO and ORB-SLAM," in 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), pp. 409-414, 2021.
[90] W. Ou, J. Zhang, K. Peng, K. Yang, G. Jaworek, K. Müller, and R. Stiefelhagen, "Indoor Navigation Assistance for Visually Impaired People via Dynamic SLAM and Panoptic Segmentation with an RGB-D Sensor," arXiv preprint arXiv:2204.01154, 2022.
[91] L. Jin, H. Zhang, and C. Ye, "A Wearable Robotic Device for Assistive Navigation and Object Manipulation," in 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 765-770, 2021.
[92] J. Xu, H. Xia, Y. Liu, and Z. Li, "Multi-functional Smart E-Glasses for Vision-Based Indoor Navigation," in 2021 6th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM), pp. 267-272, 2021.
[93] J.-L. Lu, H. Osone, A. Shitara, R. Iijima, B. Ryskeldiev, S. Sarcar, and Y. Ochiai, "Personalized Navigation that Links Speaker's Ambiguous Descriptions to Indoor Objects for Low Vision People," in International Conference on Human-Computer Interaction, pp. 412-423, 2021.
[94] G. Liu, T. Yu, C. Yu, H. Xu, S. Xu, C. Yang, F. Wang, H. Mi, and Y. Shi, "Tactile Compass: Enabling Visually Impaired People to Follow a Path with Continuous Directional Feedback," in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1-13, 2021.
[95] S. Kayukawa, T. Ishihara, H. Takagi, S. Morishima, and C. Asakawa, "Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians," in Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 4, no. 3, pp. 1-22, 2020.
[96] C.-H. Chen, C.-C. Wang, and S.-F. Lin, "A Navigation Aid for Blind People Based on Visual Simultaneous Localization and Mapping," in 2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan), pp. 1-2, 2020.
[97] H. Chen, Y. Zhang, K. Yang, M. Martinez, K. Müller, and R. Stiefelhagen, "Can We Unify Perception and Localization in Assisted Navigation? An Indoor Semantic Visual Positioning System for Visually Impaired People," in Computers Helping People with Special Needs: 17th International Conference, ICCHP 2020, Lecco, Italy, September 9–11, 2020, Proceedings, Part I 17, pp. 97-104, 2020.
[98] G. Fusco and J. M. Coughlan, "Indoor Localization for Visually Impaired Travelers Using Computer Vision on a Smartphone," in Proceedings of the 17th International Web for All Conference, pp. 1-11, 2020.
[99] H. Zhang and C. Ye, "Human-Robot Interaction for Assisted Wayfinding of a Robotic Navigation Aid for the Blind," in 2019 12th International Conference on Human System Interaction (HSI), pp. 137-142, 2019.
[100] Y. Zhao, R. Huang, and B. Hu, "A Multi-Sensor Fusion System for Improving Indoor Mobility of the Visually Impaired," in 2019 Chinese Automation Congress (CAC), pp. 2950-2955, 2019.
[101] M. Weiss, S. Chamorro, R. Girgis, M. Luck, S. E. Kahou, J. P. Cohen, D. Nowrouzezahrai, D. Precup, F. Golemo, and C. Pal, "Navigation Agents for the Visually Impaired: A Sidewalk Simulator and Experiments," in Conference on Robot Learning, pp. 1314-1327, 2020.
[102] K. Ramesh, S. N. Nagananda, H. Ramasangu, and R. Deshpande, "Real-time Localization and Navigation in an Indoor Environment Using Monocular Camera for Visually Impaired," in 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), pp. 122-128, 2018.
[103] J. Eden, T. Kawchak, and V. Narayanan, "Indoor Navigation Using Text Extraction," in 2018 IEEE International Workshop on Signal Processing Systems (SiPS), pp. 112-117, 2018.
[104] R. S. Mulky, S. Koganti, S. Shahi, and K. Liu, "Autonomous Scooter Navigation for People with Mobility Challenges," in 2018 IEEE International Conference on Cognitive Computing (ICCC), pp. 87-90, 2018.
[105] M. Lalonde, P.-L. St-Charles, D. Loupias, C. Chapdelaine, and S. Foucher, "Localizing People in Crosswalks using Visual Odometry: Preliminary Results," in ICPRAM, pp. 482-487, 2018.
[106] H. Zhang and C. Ye, "Plane-aided visual-inertial odometry for pose estimation of a 3D camera based indoor blind navigation," in 28th British Machine Vision Conference, 2017.
[107] Q. Chen, M. Khan, C. Tsangouri, C. Yang, B. Li, J. Xiao, and Z. Zhu, "CCNY smart cane," in 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pp. 1246-1251, 2017.
[108] J. Bai, D. Liu, G. Su, and Z. Fu, "A cloud and vision-based navigation system used for blind people," in Proceedings of the 2017 International Conference on Artificial Intelligence, Automation and Control Technologies, pp. 1-6, 2017.
[109] Y. Endo, K. Sato, A. Yamashita, and K. Matsubayashi, "Indoor positioning and obstacle detection for visually impaired navigation system based on LSD-SLAM," in 2017 International Conference on Biometrics and Kansei Engineering (ICBAKE), pp. 158-162, 2017.
[110] Z. Yang, L. Yang, L. Kong, A. Wei, J. Leaman, J. Brooks, and B. Li, "SeeWay: Vision-Language Assistive Navigation for the Visually Impaired," in 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 52-58, 2022.
[111] S. Shahani and N. Gupta, "The Methods of Visually Impaired Navigating and Obstacle Avoidance," in 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC), pp. 1-6, 2023.
[112] Y. Yun, T. Gwon, and D. Kim, "The Design of Person Carrier Robot using SLAM and Robust Salient Detection," in 2018 18th International Conference on Control, Automation and Systems (ICCAS), pp. 196-201, 2018.
[113] S. Sumikura, M. Shibuya, and K. Sakurada, "OpenVSLAM: A versatile visual SLAM framework," in Proceedings of the 27th ACM International Conference on Multimedia, pp. 2292-2295, 2019.
[114] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: a versatile and accurate monocular SLAM system," in IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, 2015.
[115] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras," in IEEE Transactions on Robotics, vol. 33, no. 5, pp. 1255-1262, 2017.
[116] R. Islam, H. Habibullah, and T. Hossain, "AGRI-SLAM: A Real-Time Stereo Visual SLAM for Agricultural Environment," in Autonomous Robots, pp. 1-20, 2023, Springer.
[117] T. Deng, H. Xie, J. Wang, and W. Chen, "Long-Term Visual Simultaneous Localization and Mapping: Using a Bayesian Persistence Filter-Based Global Map Prediction," in IEEE Robotics & Automation Magazine, vol. 30, no. 1, pp. 36-49, 2023, IEEE.
[118] A. Rosinol, M. Abate, Y. Chang, and L. Carlone, "Kimera: an open-source library for real-time metric-semantic localization and mapping," in 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 1689-1696, 2020, IEEE.
[119] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: Lessons learned from the 2015 mscoco image captioning challenge," *IEEE Trans-*

actions on Pattern Analysis and Machine Intelligence, vol. 39, no. 4, pp. 652-663, 2016.
[120] F. Endres, J. Hess, J. Sturm, D. Cremers, and W. Burgard, "3-D mapping with an RGB-D camera," *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 177-187, 2013.
[121] M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, and others, "Fast-SLAM: A factored solution to the simultaneous localization and mapping problem," *Aaai/iaai*, vol. 593598, 2002.
[122] X. J. Liu and Y. Fang, "Virtual touch: computer vision augmented touch-free scene exploration for the blind or visually impaired," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1708–1717, 2021.
[123] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhutdinov, R. Zemel, and Y. Bengio, "Neural image caption generation with visual attention," *Proc. ICML*, vol. 37, pp. 2048–2057, 2015.
[124] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, 2012.
[125] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 1520–1528, 2015.
[126] L. Neumann and J. Matas, "Real-time lexicon-free scene text localization and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 9, pp. 1872–1885, 2015.
[127] O. Akbani, A. Gokrani, M. Quresh, F. M. Khan, S. I. Behlim, and T. Q. Syed, "Character recognition in natural scene images," *2015 Int. Conf. Inf. Commun. Technol. (ICICT)*, pp. 1–6, 2015.
[128] K.-Y. Shao, Y. Gao, N. Wang, H.-Y. Zhang, F. Li, and W.-C. Li, "Paper money number recognition based on intersection change," *Third Int. Workshop Adv. Comput. Intell.*, pp. 533–536, 2010.
[129] M. Weber, P. Wolf, and J. M. Zöllner, "DeepTLR: A single deep convolutional network for detection and classification of traffic lights," *2016 IEEE Intell. Veh. Symp. (IV)*, pp. 342–348, 2016.
[130] J. Cheng, L. Zhang, Q. Chen, X. Hu, and J. Cai, "A review of visual SLAM methods for autonomous driving vehicles," *Engineering Applications of Artificial Intelligence*, vol. 114, p. 104992, 2022.
[131] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to autonomous mobile robots*. MIT press, 2011.
[132] L. Jinyu, Y. Bangbang, C. Danpeng, W. Nan, Z. Guofeng, and B. Hujun, "Survey and evaluation of monocular visual-inertial SLAM algorithms for augmented reality," *Virtual Reality & Intelligent Hardware*, vol. 1, no. 4, pp. 386–410, 2019.
[133] S. Zhang, S. Zhao, D. An, J. Liu, H. Wang, Y. Feng, D. Li, and R. Zhao, "Visual SLAM for underwater vehicles: A survey," *Computer Science Review*, vol. 46, p. 100510, 2022.
[134] G. Cioffi, L. Bauersfeld, and D. Scaramuzza, "Hdvio: Improving localization and disturbance estimation with hybrid dynamics vio," *arXiv preprint arXiv:2306.11429*, 2023.
[135] J. Xu, Y. Niu, and P. Shi, "Adaptive multi-input super twisting control for a quadrotor: singular perturbation approach," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 5, pp. 5195-5204, 2024.
[136] J. Xu, Y. Niu, and H. Lam, "Adaptive Distributed Attitude Consensus of a Heterogeneous Multiagent Quadrotor System: Singular Perturbation Approach," IEEE Transactions on Aerospace and Electronic Systems, vol. 59, no. 6, pp. 9722-9732, 2023.

MARZIYEH BAMDAD is currently pursuing her Ph.D. in computer science at the University of Zurich, supervised by Prof. Davide Scaramuzza, and serves as a research assistant at the Zurich University of Applied Sciences in Switzerland. Her Ph.D. research is dedicated to developing innovative solutions for visually impaired navigation, harnessing the potential of Visual Simultaneous Localization and Mapping (Visual SLAM) technologies.

DAVIDE SCARAMUZZA is a Professor of Robotics and Perception at the University of Zurich. He did his Ph.D. at ETH Zurich, a postdoc at the University of Pennsylvania, and was a visiting professor at Stanford University. His research focuses on autonomous, agile microdrone navigation using standard and event-based cameras. He pioneered autonomous, vision-based navigation of drones, which inspired the navigation algorithm of the NASA Mars helicopter and many drone companies. He contributed significantly to visual-inertial state estimation, vision-based agile navigation of micro-drones, and low-latency, robust perception with event cameras, which were transferred to many products, from drones to automobiles, cameras, AR/VR headsets, and mobile devices. In 2022, his team demonstrated that an AI-controlled, vision-based drone could outperform the world champions of drone racing, a result that was published in Nature. He is a consultant for the United Nations on disaster response and disarmament. He has won many awards, including an IEEE Technical Field Award, the IEEE Robotics and Automation Society Early Career Award, a European Research Council Consolidator Grant, a Google Research Award, two NASA TechBrief Awards, and many paper awards. In 2015, he co-founded Zurich-Eye, today Meta Zurich, which developed the world-leading virtual-reality headset Meta Quest. In 2020, he co-founded SUIND, which builds autonomous drones for precision agriculture. Many aspects of his research have been featured in the media, such as The New York Times, The Economist, and Forbes.

ALIREZA DARVISHY is professor for ICT Accessibility and head of the ICT Accessibility Lab at Zurich University of Applied Sciences in Switzerland. He serves an independent reviewer for European research projects such as the Active Assisted Living (AAL) program, and is principle investigator of the "Accessible Scientific PDFs for All" project, funded by the Swiss National Science Foundation.

• • •