Multimedia Tools and Applications (2024) 83:52171–52195
https://doi.org/10.1007/s11042-023-17552-7

Review of sensor-driven assistive device technologies for enhancing navigation for the visually impaired

Iftekar Patel[1] · **Makarand Kulkarni**[1] · **Ninad Mehendale**[1]

Received: 27 June 2022 / Revised: 29 August 2023 / Accepted: 16 October 2023 /
Published online: 9 November 2023

Abstract
Most of our daily activities hinge on our ability to perceive our surroundings, with our eyes serving as pivotal sensory receptors. Our brain processes the visual data captured by our eyes, crafting a three-dimensional representation of our environment that empowers us to navigate and execute tasks. However, individuals dealing with visual impairments experience either partial or complete absence of this visual perception. In the past, conditions like myopia and astigmatism also lacked effective treatments, yet advancements in science and technology have enabled the creation of corrective aids like eyeglasses or even minor surgical interventions for vision enhancement. Over the years, researchers have conducted extensive investigations to devise devices catering to the needs of the visually impaired. This manuscript focuses on diverse sensors applicable in navigation systems tailored for this demographic. The review delves into recent advancements in navigation tools designed for the visually impaired, encompassing sensors such as visual, proximity, and LiDAR sensors, among others. Sensors generate copious amounts of data, which undergo processing to simulate the surrounding environment. We underscore the unique capabilities of each sensor, as well as optimal combinations of sensors that yield superior results. The challenges associated with sensor utilization are also addressed, accompanied by potential strategies for overcoming them. Our survey reveals a prevalent trend in utilizing RGB-depth sensing cameras and ultrasonic sensors in tandem with other sensor types for navigation purposes. This study aims to furnish a comprehensive overview of contemporary progress within the realm of navigation aids for the visually impaired, ultimately aiding researchers in discerning the most suitable technological approaches based on specific applications.

Keywords Navigation systems · Obstacle detection · Visually impaired · RGB depth sensing cameras · LiDAR navigation

✉ Ninad Mehendale
ninad@somaiya.edu

[1] K. J. Somaiya College of Engineering, Somaiya Vidyavihar University, Vidyavihar, Mumbai, India

1 Introduction

Visual impairment encompasses both complete and partial loss of eyesight, often involving conditions that cannot be corrected through the use of corrective eyewear. Globally, approximately 270 million individuals grapple with visual impairments [4]. The origins of visual impairment vary and can stem from factors such as childhood blindness, advanced age, infections, or brain trauma. While many individuals blind from birth can achieve educational and professional excellence through tailored training, certain impairments manifest later in life due to aging or infections. Notably, cataracts stand as a prevalent cause of blindness among humans. Despite the highly successful outcomes of cataract surgeries, medical practitioners must carefully consider the patient's age and comorbidities. An additional contributor to blindness is cortical visual impairment [13], which arises when a person's brain sustains damage from occurrences like strokes or accidents.

Technology has made tremendous strides in recent years to address and aid visual impairments. While rectification remains unattainable for some individuals, they resort to utilizing assistive tools like canes for mobility support. However, this reliance on canes not only occupies the person's hands but also restricts their ability to have their hands free.

Numerous researchers have conducted studies aiming to create hands-free navigation solutions for the visually impaired. Sensors like ultrasonic sensors [14], vision sensors [24], Light Detection and Ranging (LiDAR) sensors [9], and Infrared (IR) sensors [17] have been harnessed for the development of assistive devices. Sensors find wide utility in scenarios necessitating precise detection of motion, obstacles, proximity, temperature, and other environmental fluctuations [42]. These sensors intake environmental changes and produce outputs intelligible to microcontrollers and computers. They find applications across diverse domains, opening up endless possibilities when integrated into various systems. For tasks demanding motion and obstacle detection around individuals, non-contact sensors come into play. Non-contact sensors emit sound or light beams and, upon encountering an obstacle, these beams reflect back to the sensor [39].

Figure 4 shows sensors used in navigation systems arranged in a hierarchical order. Sensors such as ultrasonic , Infrared (IR), LiDAR, accelerometer, Inertial Measurement Unit (IMU), and Global Positioning System (GPS) have been widely used in the development of navigation systems. Depth-sensing (RGB-Depth) cameras, simple RGB cameras, and stereo cameras are the three types of visual sensors (cameras). Then, based on their coverage area, LiDARs are classified as 1D, 2D, and 3D. Some investigations used IMU in conjunction with LiDARs and cameras, to improve the pose detection of the user. Further, to improve obstacle detection accuracy, the 1D LiDAR was employed with the RGB and RGB-Depth (RGB-D) cameras. GPS is also used in conjunction with other sensors to determine a person's location in relation to a map and then navigate them in the correct direction. The scope of this manuscript includes the different sensors that are commonly used for obstacle detection while an emphasis has been given on LiDAR sensors. The manuscript also discusses what sensors used individually or combined with other sensors will be the best for navigation.

2 Infrared sensors

IR sensors find widespread usage in scenarios where the detection of obstacles in a straight line is required. Typically affixed at specific points on navigation systems, these sensors are positioned to promptly identify obstacles directly ahead or discern recesses in the ground. The placement of the IR sensor and its distance from obstacles significantly influence navigation accuracy.

In the work by Innet et al. [17], a walking stick equipped with an IR sensor was developed to identify various obstacles situated in front of the user. The primary objective was to ascertain obstacles within an 80 cm range while also detecting diverse materials such as metals, glass, wood, concrete, and even human entities.

$$Output\ Voltage = \frac{4187.8 \times range^{-0.9042}}{208.8555} \tag{1}$$

Equation (1) demonstrates the computation of the returning IR beam's value by the microprocessor. This voltage value is directly proportional to the obstacle's distance. The researchers utilized a PIC16F877 microcontroller for their system. Their setup provided distinct alerts contingent on the obstacle's proximity. They achieved accurate obstacle detection within the range of 10 to 55 cm, although certain materials remained undetected beyond that threshold.

Nada et al. [33] developed a walking cane for the visually impaired, equipping it with an IR sensor affixed to its base. This enhancement facilitated the detection of depressions in the ground, such as staircases, while also enabling the identification of other obstacles within a 2-meter range. Their system employed a PIC16f877 microcontroller and utilized audio messages to provide user alerts. The system exhibited avoidance accuracy ranging from 75

These sensors find frequent application in walking canes and obstacle detection devices. However, their detection range is limited, and accuracy diminishes as the distance increases. While these sensors excel at short-range obstacle detection, they may not be the optimal choice for navigation purposes alone. Nonetheless, they can be effectively integrated into navigation systems alongside other sensors to enhance obstacle perception within the environment.

3 Ultrasonic sensors

Gbenga et al. [15] employed ultrasonic sensors and Arduino to create a walking stick. Ultrasonic sensors were strategically positioned on the stick and interfaced with an Arduino board. When an obstacle was detected by the sensor, a buzzer was activated, signaling the user. The effective sensing range of their stick spanned from 0 to 2 meters. Moreover, their system showcased the ability to identify water presence through a moisture sensor.

Frizziero et al. [14] investigated various systems catering to the visually impaired that incorporated ultrasonic sensors and feedback mechanisms. They adopted methodologies such as Design for Six Sigma (DFSS), coupled with Quality Function Deployment (QFD) and Stylistic Design Engineering (SDE), to establish a comprehensive workflow for analyzing diverse systems. Their evaluation encompassed parameters like safety, simplicity, detection range, weight, cost, ergonomic design, battery life, charging duration, and waterproofing.

4 Vision sensors

RGB camera

Queueing is a routine task for many individuals, yet it presents challenges for those with visual impairments who struggle to navigate and advance within the line independently. To address this, Kuribayashi et al. [24] devised a solution enabling visually impaired individuals to navigate queues autonomously. Their system utilizes an off-the-shelf iPhone 11 Pro equipped with a rear-end RGB camera, an infrared depth sensor, and an IMU sensor. Augmented Reality (AR) markers placed around the room, along with a pre-existing floor map, facilitate the user's localization within the space. By scanning AR markers, the system identifies the user's position, providing navigation instructions to guide them to the queue's end. YOLOv3-tiny processes the smartphone's rear camera video stream to detect people and delineate bounding boxes around them. The infrared depth sensor calculates the distances to the identified individuals, designating the nearest person as the target. The user is guided to stand behind the target at a safe distance, factoring in social distancing guidelines. The RGB camera and depth sensor continuously monitor and track the target, employing distinct vibrations to convey various messages. Prolonged vibration pulses prompt the user to halt, while dual-pulse vibrations signal them to move forward. Short vibrations alert the user to obstacles ahead. The system underwent testing with 12 visually impaired participants, achieving an average System Usability Scale (SUS) score of 83.9 out of 100.

RGB-Depth camera

An RGB-depth camera captures object distances within an image, representing these distances through distinct colors. The pixel intensity values enable distance calculation for captured objects, and object location can be inferred based on pixel detection of nearby objects. Figure 1(f) illustrates how a depth-sensing camera captures depth imagery, where various color zones correlate to object distances from the camera, transitioning from blue for close objects to red for distant ones.

Mostofa et al. [32] developed a walking aid for the visually impaired, exploring sensors like ultrasonic, RGB-D, and basic RGB cameras. They evaluated haptic and audio feedback systems, incorporating 9 ultrasonic sensors on the walker's lower section. However, they observed signal interference and inaccurate readings due to sensor proximity, leading them to devise a sequential algorithm for individualized sensor readings. These ultrasonic sensors detected within a 200-meter range. The team then employed an RGB camera alongside computer vision, utilizing Faster R-CNN with Inception-ResNet V4 to train a model on the Open Images dataset. Distance estimation relied on object area within an image. Microsoft Kinect's RGB-D camera located obstacles, employing direct depth image processing and point cloud generation, the latter being lightweight in computation. Figure 1(c), (d), and (e) depict their point cloud approach. After testing configurations including a sensor-less cane, a cane with ultrasonic sensors, and a cane with visual sensors, the camera-ultrasonic hybrid configuration exhibited accurate and swift navigation.

Zhang et al. [50] developed a navigation aid for the visually impaired utilizing an RGB-D camera. Their system employs a 6 Degree of Freedom (DoF) robot to gather both depth and non-depth data from the cameras, as well as IMU values, to ascertain the robot's position and orientation, thus facilitating navigation. To mitigate potential visual data errors, they implemented a particle filter localization method, leveraging information from a developed 3D CAD model of indoor spaces. The proposed system achieved a remarkable 82.5% reduction in pose error, while maintaining an 18 Hz update rate for pose estimation.

Hakim et al. [16] devised a navigation framework employing an RGB-D camera coupled with an ultrasonic sensor for mapping, localization, and obstacle detection. The Asus Xtion Pro RGB-D camera generates a 2D image of the surroundings, which is subsequently processed through Hector SLAM to construct a map and determine the user's location within it. Operating within a distance detection range of 0.8 to 3.5 meters, the RGB-D camera facilitates path planning by generating a 2D map for identifying start and end points. The A*

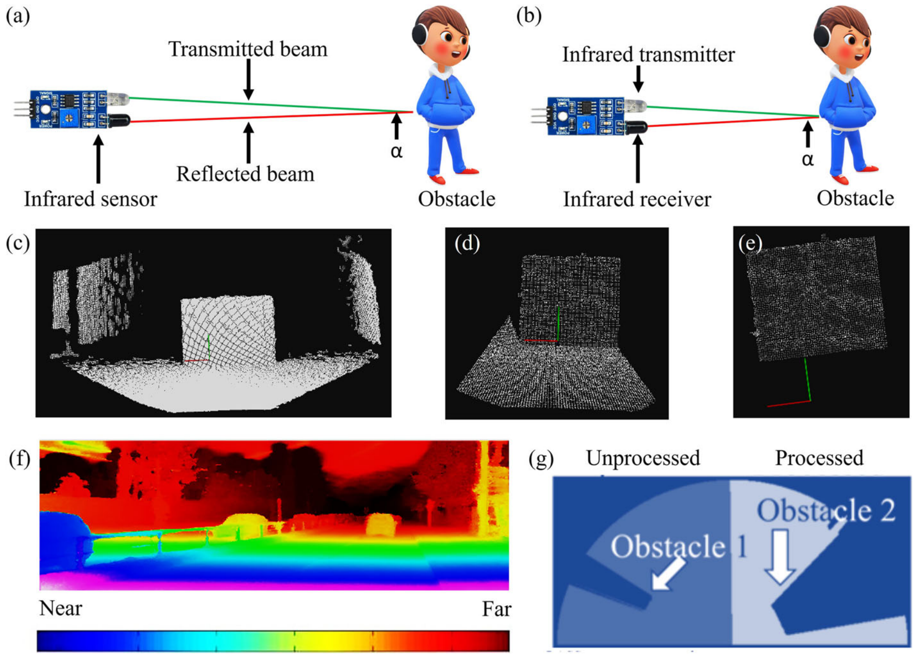

Fig. 1 (a) Working of the IR sensor as stated by Nada et al. [33]. The obstacle is far away from the IR sensor and hence the angle α created between the transmitted and reflected rays is small. (b) The obstacle is nearer to the IR sensor and hence the angle α created between the transmitted and reflected rays is larger. Using this the distance of the obstacle from the visually impaired person was estimated [33]. (c) Point cloud processing approach proposed by Mostofa et al. [32]. The point cloud of the surrounding environment was captured directly by the Kinect camera. (d) Points that were no longer relevant were removed. (e) After the floor was discovered, those points were eliminated, leaving only the obstacle's point cloud. (f) The depth image is captured by a depth-sensing camera. The different color zones indicate the distance of the objects in the image when seen from the camera. The color spectrum begins at blue pertaining to near objects and red pertaining to far objects. (g) The operation of the LASS system as shown by Ton et al. [44]. Obstacle 2 is recognized first, and the user is notified, followed by the alert for obstacle 2. Obstacle 2 is also processed before Obstacle 1 because scanning in that location occurs first. Similarly, because it is closer to the user, obstacle 2 has a higher pitch

algorithm then computes an optimal path. The ultrasonic sensor serves to detect obstacles along the designated path, with simple RGB images, lacking depth data, used for object identification. In their study, they adopted a visually impaired person's step length of 35 cm. Through testing in various environments, they established the system's reliability in guiding individuals from one point to another.

Stereo camera

Numerous solutions are currently in development to aid individuals with visual impairments in navigating their surroundings. However, these individuals require assistance beyond navigation, encompassing tasks like object identification, reading signs, and recognizing currency denominations. To tackle this multifaceted challenge, Rahman et al. [37] have devised a comprehensive system capable of furnishing both navigation guidance and support for daily activities. The system integrates ultrasonic sensors with a stereo camera setup, engaging user interaction through auditory commands. Ultrasonic sensors gauge object distances ahead of the user, while the stereo camera captures images, subsequently analyzed by the YOLO object detection algorithm. Upon user command, the stereo camera captures the scene, YOLO identifies objects, and audio alerts relay the information to the user. This empowers the user to select a desired object for navigation, receiving directional guidance from the system. The system further extends its capabilities to recognize familiar faces, notifying the user upon detection. By amalgamating ultrasonic distance estimation with YOLO's object detection, precise object distances can be calculated. Additionally, Optical Character Recognition (OCR) plays a crucial role, detecting, processing, and audibly conveying text within images captured by the camera. The system adeptly identifies text in Bengali and reads currency denominations aloud. The reported object detection accuracy stands at an impressive 98.9%. The integration of a Global Positioning System (GPS) module offers real-time user location tracking, facilitated through a dedicated mobile app or online map platform. Notably, the system's speech recognition algorithm enables seamless user-device interaction, capturing and processing spoken input. Moreover, a feature for immediate emergency communication is provided, enhancing the system's utility for visually impaired users.

Toro et al. [45] developed a navigation system designed to guide visually impaired individuals through unfamiliar environments by detecting obstacles, open spaces, and objects such as doors, chairs, computers, and more. Upon detection, they employ a path planning algorithm that guides the user to their intended destination. The system is divided into six distinct components. The first component involves floor segmentation, utilizing depth data from the camera and rotating the obtained point cloud based on the camera's position and orientation, determined by the camera's IMU. The second component creates an occupancy grid, storing information about occupied and free pixels in the camera's captured image at a given pose. The third step involves obstacle detection, which transforms the grid into a distance map. The fourth component identifies objects of interest, utilizing the You Only Look Once (YOLO) model to detect objects and locate their position relative to the user using depth data. Finally, haptic feedback is delivered through a vibrator motor attached to a belt. The system employs a GPU for real-time processing of depth and RGB data from the stereo camera, and the entire setup is wearable. Additionally, the user receives audio feedback for alerts.

Rahman et al. [37] have devised a system incorporating ultrasonic sensors for obstacle detection and a stereo camera to gain insights into the surrounding environment. The camera captures images of the surroundings, subsequently subjecting them to the YOLO V4 deep learning model for object identification. Furthermore, they employ a face recognition algorithm, employing feature landmark estimation to recognize individuals and perform optical character recognition.

In a distinct approach, Vorapatratorn et al. [46] have developed a navigation assistance system tailored for the visually impaired, relying solely on a stereo camera setup. The images captured from the left and right perspectives of the stereo setup undergo processing to estimate the depth of each pixel. This depth information is then translated into Horizontal-Depth Accumulative Information (H-DAI), streamlining the elimination of the ground plane and enhancing the visibility of regions containing obstacles. The highlighted structures, as delineated by H-DAI, are projected onto a bird's-eye view plane using Vertical-Depth Accumulative Information (V-DAI). The ensuing depiction of obstacles, post V-DAI processing, is fed as input to various supervised machine learning models such as Artificial Neural Network (ANN), Support Vector Machine (SVM), and decision trees. These models are trained to analyze the bird's-eye view representation of obstacles and categorize them into seven distinct obstacle classes. The output of their system is depicted in Fig. 5(e) (f) (g). The training dataset encompasses 34,325 grayscale stereo images, each meticulously annotated for object identification. Once all detected objects are classified, navigation instructions to circumvent them are relayed to the user via Bluetooth headphones. Impressively, they achieve an object classification accuracy of 96.45% at a processing speed of 23.76 images per second.

5 LiDAR

5.1 Different uses of LiDARs

Zhou et al. [52] proposed a 3D object detection network that eliminates the need for prior feature engineering. Many existing 3D object detection algorithms relied on manual feature engineering applied to raw LiDAR point cloud data. VoxelNet, however, takes raw LiDAR point cloud data as input, detects and classifies 3D objects into categories like cars and pedestrians, and outlines bounding boxes around each detected object. VoxelNet comprises three main components. The first is the Feature Learning Network (FLN), which divides the LiDAR point cloud data into uniform 3D boxes known as voxels. Each LiDAR data point within a voxel is transformed into a vector, maintaining the shape of the data points while creating a sparse 4D tensor. The second section features convolutional layers that extract essential information from the 4D tensors. The extracted data then flows into a Region Proposal Network (RPN), producing a probability score map and a regression map for detecting and categorizing 3D objects within an input LiDAR point cloud.

Initially, they collected 4000 training point clouds from the KITTI dataset and expanded the dataset size using augmentation. The KITTI dataset contained 7481 training and 7518 testing point clouds, offering ground truth data for three types of 3D objects: cars, pedestrians, and cyclists. The dataset included two perspectives-a 3D view and a top view known as Bird's eye view. The KITTI dataset categorized LiDAR point clouds into three difficulty levels: easy, moderate, and hard. VoxelNet achieved precision accuracies of 77.47, 65.11, and 57.73 for the 'car' object class across the three difficulty levels (easy to hard) in the 3D view. For the 'pedestrian' object class in the 3D view, it achieved precision accuracies of 39.48, 33.69, and 31.51 across the same difficulty levels. Lastly, for the 'cyclist' object class in the 3D view, it attained precision accuracies of 61.22, 48.36, and 44.37 across the respective difficulty levels.

Xu et al. [47] state that LiDAR scans exhibit a 2.5D nature, where objects are detected through laser beams reflecting off them. However, critical information regarding the 3D scene situated behind these objects is lost, resulting in an incomplete perception of the entire 3D object shape, referred to as "shape miss." This phenomenon is attributed to three factors: self-occlusion, external occlusion, and signal loss. Self-occlusion occurs when certain segments of an object block laser beams from reaching other parts of the same object. External occlusion arises when an external object obstructs laser beams from accessing specific portions of another object. Signal loss transpires when transmitted laser beams fail to be received by the LiDAR sensor due to peculiar reflection angles.

To address this issue, Xu et al. [47] propose a model named the "Behind the Curtain Detector," aimed at detecting 3D objects substantially affected by occlusion. Their model initially identifies regions characterized by occlusion and signal loss. A shape occupancy network estimates the likelihood of object shape occupancy. Extracting 3D features from point cloud data is achieved through a backbone network. These features serve as inputs for the region proposal network (RPN), generating 3D proposals. The refinement of proposals incorporates geometric features derived from the multi-scale features of the backbone network and the estimated shape occupancy probability. This refinement process ultimately yields bounding boxes for each identified 3D object within the input LiDAR point cloud.

While 3D or solid-state LiDARs represent the ideal choice for navigation systems, their high cost poses a significant hurdle. As a more cost-effective alternative, the deployment of single laser beam versions, namely 2D LiDARs, becomes necessary. These 2D LiDARs offer precise distance measurements within a 360-degree horizontal field of view. Nevertheless, their limitation lies in the inability to detect objects at varying heights, both low and high. To surmount this challenge and amalgamate these methods effectively, the integration of cameras emerges as a promising solution. However, cameras fall short in providing accurate distance measurements. To address these shortcomings, Young et al. [49] present an innovative approach that combines LiDAR and cameras. They introduce a technique segmenting obstacles into three categories: in line, low, and high. The LiDAR sensor detects objects directly ahead, with the camera activated upon LiDAR detection to identify hanging obstacles. This dynamic switching conserves energy, while their SIFT algorithm enables faster obstacle detection without interrupting sensor operation. The study highlights SIFT's efficacy in detecting hanging obstacles and underscores the advantages of employing solid-state drives (SSDs) for memory storage, given their resilience to vibrations and enhanced read/write speeds that boost mapping and detection accuracy.

Cao et al. [7], on the other hand, propose a trajectory planner designed to mitigate LiDAR sensor detection failures. Their method formulates a path based on both detected objects and raw sensor data, utilizing a CNN-Segmentation neural network for object detection. The trajectory planner incorporates multiple layers to accommodate the varied perceptual inputs. To prevent collisions stemming from detection gaps, the system dynamically adjusts attention to either detected objects or point cloud data. Results demonstrate that the system efficiently plans trajectories even when nearby objects remain undetected, ensuring the autonomous vehicle's safe navigation. Simultaneously, when the perception system functions optimally, the vehicle remains unaffected by point cloud variations. This strategy bolsters the reliability of autonomous vehicles and enhances their trustworthiness.

PanoVILD, developed by Javed et al. [20], presents a challenging dataset that combines panoramic vision and LiDAR data collected by an autonomous vehicle. Equipped with a Point Grey Ladybug 3 camera, 3D LiDAR, GPS, and an inertial measurement unit (IMU), the vehicle gathers data during outdoor drives encompassing diverse scenes, including a parking lot, a semi off road path, and a campus road with traffic. Realtime synchronization and registration of data from all vehicle mounted sensors are ensured. The dataset encompasses 3D LiDAR point clouds, images, GPS, and IMU measurements. The images hold potential for the development and validation of novel multifisheye, panoramic, and 3D LiDAR-based simultaneous localization and mapping (SLAM) systems. Designed for various applications such as odometry, SLAM, loop closure detection, and deep learning-based algorithms involving vision, inertial, LiDAR, and fusion of visual, inertial, and 3D data.

Notably, 2D Fully Convolutional Networks (FCN) have demonstrated remarkable object detection accuracy in images. This capability has been extended to detecting 3D objects within point cloud data by Li. The object detection process using FCNs comprises two stages: objectness prediction and bounding box prediction, both culminating in output maps that constitute the FCN. Objectness maps discern the presence of objects in specific areas, while bounding box maps generate coordinates for bounding boxes. Raw point cloud data is processed by passing it through three convolutional layers with a $1/2^3$ stride to downsample, followed by upsampling with a convolutional layer of matching stride to yield objectness and bounding box maps. Their proposed VeloFCN system underwent training and evaluation using the KITTI dataset, yielding Average Precision (AP) values of 93.7%, 81.9%, and 79.2% for the easy, moderate, and hard sections of the image plane.

An integral facet of autonomous driving lies in the capability to recognize pedestrians. Equipped with an array of sensors such as LiDAR and cameras, autonomous vehicles can process data from each of these devices concerning their surroundings. Premebida et al. [36] proposed a pedestrian detection system utilizing RGB values from cameras and 3D point clouds from a LiDAR sensor. The study also explored the upsampling of depth maps from the LiDAR sensor. Two strategies were advanced to efficiently combine RGB images with point cloud data for pedestrian detection. The KITTI dataset was employed for model training and assessment, with performance compared to the deformable parts model serving as the baseline detector. Their first fusion approach involved crafting a combined vector from the feature vectors extracted from both RGB images and point cloud data, serving as input for a detector. This method, referred to as centralized fusion, entailed feature-level fusion. The second approach comprised training separate detectors for image and depth data, subsequently fusing the detectors' outputs for the final result. This decentralized fusion strategy operated at a higher decision level, and it exhibited greater efficacy in pedestrian detection, achieving average precision values of 59.86%, 49.38%, and 43.18% for the KITTI dataset's easy, moderate, and hard categories, respectively.

Extensive efforts have been dedicated to 3D object detection based on image and point cloud data. Nevertheless, Chen et al. [11] harnessed front and bird's eye views derived from the LiDAR sensor, coupled with RGB data, to generate oriented 3D bounding boxes. A multi-view encoding scheme was initially employed to generate a comprehensive view of the sparse 3D point cloud from bird's eye, front, and RGB perspectives. Their system encompassed two components: a 3D proposal network and a region-based fusion network. The bird's eye view of the LiDAR sensor facilitated accurate 3D object proposals via the 3D proposal network. Subsequently, the region-based fusion network projected these 3D object proposals onto each of the three input views for object detection. Performance evaluation was conducted on the KITTI dataset, yielding Average Precision (AP) values of 96.52%, 89.56%, and 88.94% for the dataset's easy, moderate, and hard categories, respectively, with an Intersection-over-Union (IoU) value of 0.25 for the 3D object detection task.

5.2 LiDAR-based navigation systems for the visually impaired

1D LiDAR

Detecting ground-level obstacles, such as staircases, stands as a crucial task for navigation systems aimed at assisting the visually impaired. Chai et al. [9] have developed a LiDAR-based ground plane assessment system capable of identifying staircases, ramps, and potholes among other obstacles. Details about these obstacles are transmitted to the user through a mobile app, as depicted in Fig. 2. Their design encompasses a band housing a micro 1D LiDAR sensor, a servo motor, a Bluetooth module, and a Micro Arduino computing unit. The LiDAR sensor is inclined at a fixed angle, and the band is worn near the knee. Monitoring the LiDAR sensor's distance readings is a constant process. If readings increase or decrease beyond a specific threshold, the servo motor tilts the LiDAR sensor upwards. By analyzing distance readings at both angles, their system discerns upward or downward stairs, ramps, and potholes. This information is wirelessly conveyed to the mobile app via Bluetooth. The system underwent testing across ten distinct paths, including ramps, staircases, and potholes, achieving an impressive overall accuracy rate of 93.1%. However, due to the usage of a 1D LiDAR sensor, the system's capability to precisely identify other ground-level obstacles, such as gravel, remains limited.

In a parallel endeavor, Bouteraa et al. [5] devised a LiDAR and fuzzy logic-driven system to enhance indoor navigation for visually impaired users. This multifaceted system incorporates a 1D LiDAR sensor, ultrasonic sensors, an IMU sensor, a Beagle Bone Black microcontroller, and Bluetooth speakers. Ultrasonic and IMU sensors find placement on a pair of eyeglasses, while the 1D LiDAR resides on a glove worn by the user. The system continuously monitors user speed and obstacle distances, which serve as inputs for a fuzzy logic-based classifier. This classifier employs a rule table to determine the appropriate safety assessment based on user speed and obstacle proximity. Once the classifier ascertains the safety level, audio alerts are transmitted to the user via Bluetooth headphones. The system underwent trials with 7 participants, demonstrating significantly reduced navigation times through unfamiliar paths compared to traditional walking aids.

Concurrently, Chitra et al. [12] devised an audio-enabled navigation device for the visually impaired. This device features a 1D LiDAR sensor and camera affixed to a waist belt for navigation purposes. The LiDAR calculates obstacle distances, while the camera captures frames used for object identification. Object classification is facilitated by a Convolutional

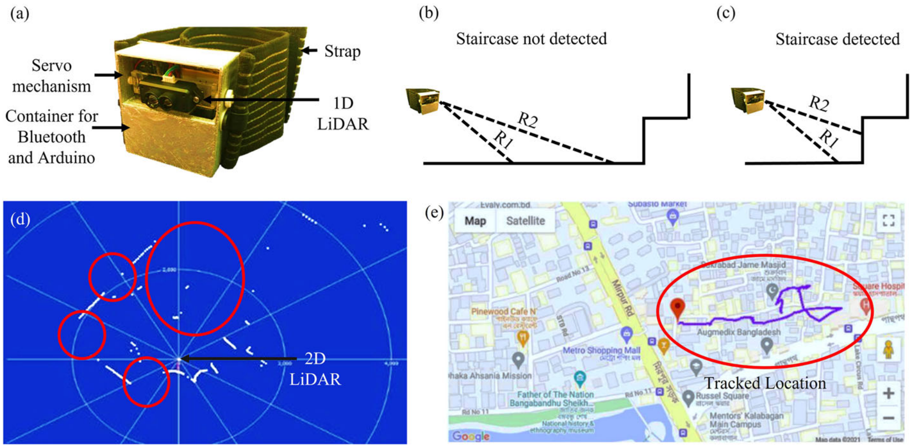

Fig. 2 Ground-based obstacle detection system proposed by Chai et al. [9] to detect staircases, ramps, potholes, etc. (a) Components of the proposed system consist of a 1D LiDAR sensor, servo motors, Bluetooth module, and an Arduino microcontroller. The device is to be worn above the knee. (b) Case: no obstacles. The LiDAR sensor is initially positioned at a fixed downwards angle. The distance computed from the sensor readings is monitored to detect ground-based obstacles. When an obstacle is in front of the user the distance measured increases or decreases depending upon the nature of the obstacle. In such a case the LiDAR sensor is tilted upwards by a fixed amount using the servo mechanism. (c) Case: upwards stairs. The distance calculated at the starting angle of the LiDAR sensor is the same as before but the distance calculated at the second angle has decreased hence the system concludes that the user is facing upward stairs. This information is relayed to a mobile app via Bluetooth, which then alerts the user.(d) The obstacles as seen by the 2D LiDAR using a visualization tool [23]. The white lines depict the obstacles around the LiDAR sensor. The red circled areas show the openings as seen by the LiDAR. These openings are incorporated with ultrasonic sensor data and computer vision to navigate the user. (g) Rahman et al. [37] have provided a facility that allows the system to keep a track of the visually impaired person's location and also remlembers their past locations and visited sites. This data is incorporated with a mobile application or online map platform. The purple line depicts the user's path to the intended place

Neural Network (CNN). The entire system is managed by a Raspberry Pi module. Users receive audio alerts containing information about the obstacle type and distance, while those with hearing impairments benefit from vibrational feedback provided by an attached motor. Nevertheless, this work is confined to identifying a limited number of objects.

Chen et al. [10] developed a system that utilizes visual Simultaneous Localization and Mapping (SLAM) to establish the user's position based on pre-saved maps. As the user moves, the system continuously analyzes the pre-saved map, employing visual input features to align the user's location with the map. Their system incorporates audio feedback and integrates RGB-Depth (RGB-D) cameras, Inertial Measurement Units (IMUs), and LiDAR sensors. The IMU aids in mapping the surroundings and charting the route, while the 1D LiDAR sensor is employed to mitigate discrepancies arising from visual inputs.

2D LiDAR

The team responsible for the development of a pedestrian navigation system for visually impaired individuals [22] has also designed a system to guide users toward unoccupied outdoor seating. In their work, Kayukawa et al. [21] employed RGB depth cameras and a 2D LiDAR sensor mounted on a commercially available robot to detect vacant chairs and direct users to them. The utilization of YOLOv3 on RGB camera images enabled chair identification, while proximity of individuals to detected chairs was ascertained using YOLOv3 to determine occupancy status. Unoccupied chairs were designated as targets based on the absence of nearby individuals, with the 2D map of the scene and distance estimation to the target chair being derived from RGB depth camera and 2D LiDAR data. The path to the target chair was generated using the Robotic Operating System (ROS) navigation stack, which the robot followed while the user held onto a handle. A test involving 6 participants demonstrated that the system facilitated users' prompt access to empty chairs in a controlled environment, surpassing the efficiency of an audio-based navigation system.

Khaitan [23] initiated work on a navigation aid for the blind, evaluating various components including cameras, 2D LiDAR, and ultrasonic sensors by integrating them onto a rover. The ultrasonic sensor's 180-degree front and back views were partitioned into five segments, sending a 5-bit sequence indicating presence (1) or absence (0) of obstacles. The Pixy2 camera identified objects in images, with a focus on specific aspects rather than comprehensive frame analysis. A 2D LiDAR, capturing images within a single plane, was adapted for 3D surroundings by altering angles through a trigonometric formula (2) due to vertical LiDAR movement impacting distance calculations. Optimal implementation involves merging this approach with an image recognition algorithm. Figure 2(d) illustrates obstacles detected by the 2D LiDAR using a visualization tool, displaying obstacle positions (white lines) and LiDAR apertures (red-ringed areas). User guidance combines these apertures with ultrasonic data and computer vision.

$$\cos\theta = \frac{base}{hypotenuse} \tag{2}$$

Subsequently, a YOLO deep learning model was deployed for object perception in the surroundings, classifying 80 classes with this algorithm. Further enhancing their system's object detection capabilities, they trained a model on the CIFAR-10 dataset, achieving the highest accuracy of 80

Ton et al. [44] devised a LiDAR Assist Spatial Sensing (LASS) system to aid visually impaired individuals in navigation, addressing limitations associated with echolocation-based navigation systems. Their innovative approach involves utilizing LiDAR data to acquire spatial information, translating it into stereo sound with varying pitch levels. The pitch conveys object pose in relation to the user. Emitting a 785 nm laser beam, their LiDAR system provides a detection range of 4.1 meters. The system converts obstacle angle data into pitch and distance information into audio with distinct frequencies. Environmental data spanning a 180-degree front view is sent for processing to a computer. Employing the SoundStream class from the Simple and Fast Multimedia Library, the system generates sounds based on distance and angular data. Varying sound intensity and timing during the transition from left to right ear approximates the object angle for the user. Higher pitch indicates closer obstacles, with more pronounced increases for objects in proximity. The system scans the surroundings within a 180-degree semi-circle every 13 seconds, as illustrated in Fig. 1(g).

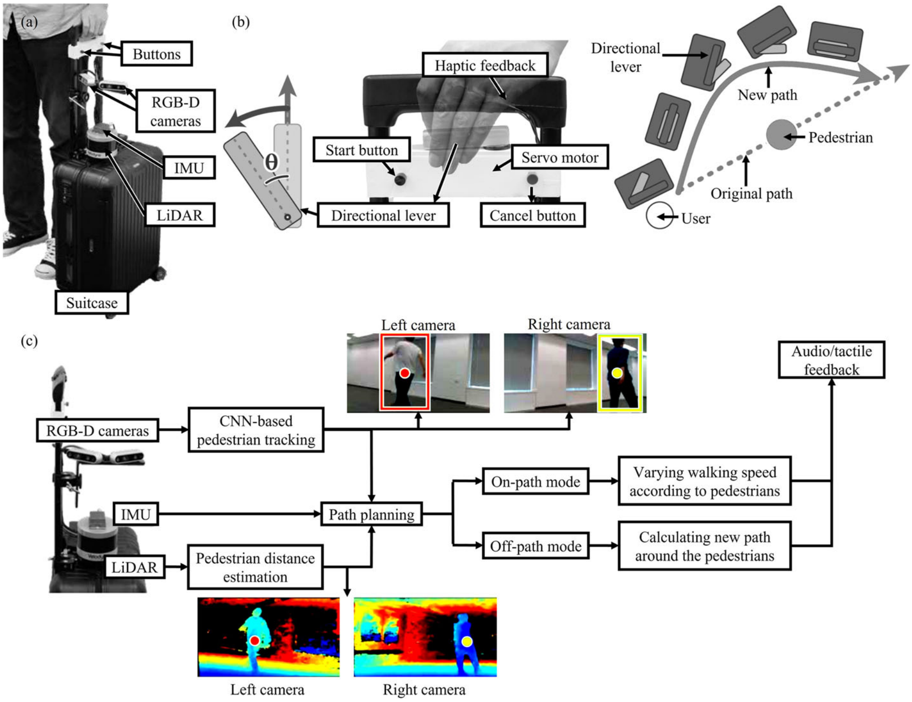

Fig. 3 Navigation system designed by Kayukawa et al. [22] for visually impaired pedestrians to help them navigate crowded areas. (a) A regular wheeled suitcase was fitted with two RGB depth cameras, an IMU sensor, a 3D LiDAR, and control buttons on the handle. (b) Feedback mechanism for the user. The handle of the suitcase consists of a vibration motor, servo motor, and a directional lever. The directional lever is rotated by using the servo motor and it indicates the direction of the path to be followed by the user to avoid collisions with other pedestrians. The vibration motor is used to provide haptic feedback to the user. (c) Pedestrians are detected and tracked via a Convolutional Neural Network (CNN) model applied on the video stream captured by the RBG cameras. The depth information captured by the cameras and the 3D point clouds recorded by the 3D LiDAR was used to calculate the position of the pedestrians detected by the CNN model. In the on-path mode of operation, the speed of the user is monitored via the IMU sensor, and the user is asked to change their walking speed in accordance to the pedestrians in front of them. In the off-path mode, a new path is planned when there are pedestrians in front of the user. Navigation instructions are presented to the user via audio/tactile feedback as shown in (b)

The sequence shows obstacle 2 being detected first, triggering user alert, followed by obstacle 1's alert. Obstacle 2's prompt precedes that of obstacle 1 due to the initial scanning focus. Moreover, the higher pitch of obstacle 2 corresponds to its closer proximity. Notably, optimal performance was observed when obstacles were at a 90-degree angle.

Maryono et al. [29] utilized a 2D LiDAR in conjunction with the K-means clustering algorithm to detect and classify obstacles within the vicinity of visually impaired individuals. The point cloud data gathered by the 2D LiDAR underwent conversion into a data array. The K-means clustering algorithm was subsequently applied to group data points belonging to a single object.

In a manner reminiscent of animals like cats using their whiskers to navigate in low-light conditions, Pallej'a et al. [34] developed an obstacle detection system for the visually impaired employing an accelerometer and a 2D LiDAR sensor. The 2D LiDAR scans a 360^o plane, analogous to the whiskers' role for animals. When objects breach this plane, the 2D LiDAR captures their shape and distance, mirroring the whisker's function. The user wears a wristband equipped with an accelerometer to detect orientation. The system employs haptic feedback, manifesting as varying pulses, to relay obstacle location and distance information, intentionally avoiding disruptive audio alerts that could interrupt daily activities. Testing demonstrated over 95% accuracy in detecting plain floors, ascending, and descending stairs, as well as 100% accuracy in identifying small-volume objects. However, detection accuracy decreased as object volume increased. The whisker concept proposed by Pallej'a et al. [34] is illustrated in Fig. 5 (a). The LiDAR sensor detects the flat floor, with the red conical shape symbolizing the whisker. Figure 5 (b) illustrates the flat floor graph, appearing as a straight horizontal line. When a cubic obstacle enters the LiDAR sensor's range, mimicking whisker contact (c), the graph reflects the newfound obstacle's presence (d). The sides correspond to the flat floor, while the indented center depicts the obstacle.

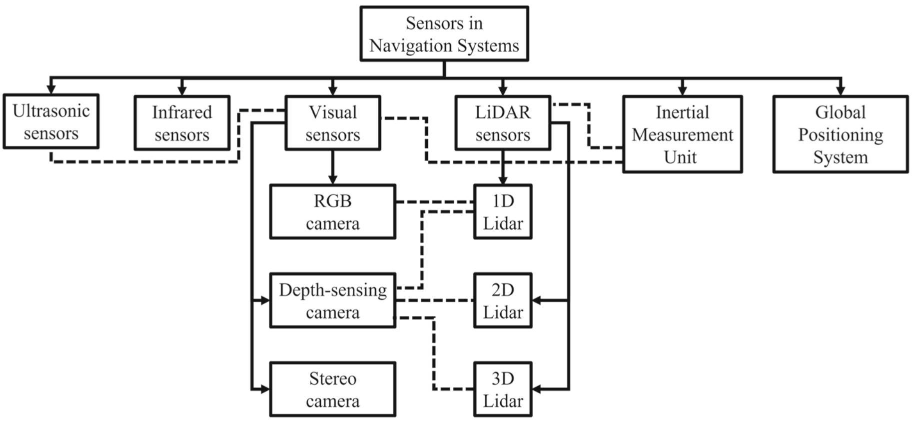

Fig. 4 The hierarchy of sensors employed in navigation systems encompasses various types. Among these, ultrasonic, infrared, LiDAR, Inertial Measurement Unit (IMU), and Global Positioning System (GPS) sensors have found frequent usage in the development of navigation systems. Visual sensors, represented by cameras, can be further categorized into depth-sensing (RGB-Depth) cameras, conventional RGB cameras, and stereo cameras. The LiDAR category is sub-divided into 1D, 2D, and 3D variants, contingent on their spatial coverage. Notably, certain studies integrate IMUs with LiDARs and cameras, illustrating their interconnected functionality. Furthermore, instances arise where 1D LiDARs are combined with RGB and RGB-Depth (RGB-D) cameras to enhance obstacle detection precision. GPS is also synergistically employed alongside other sensors, facilitating location estimation vis-à-vis a map and subsequent accurate navigation guidance

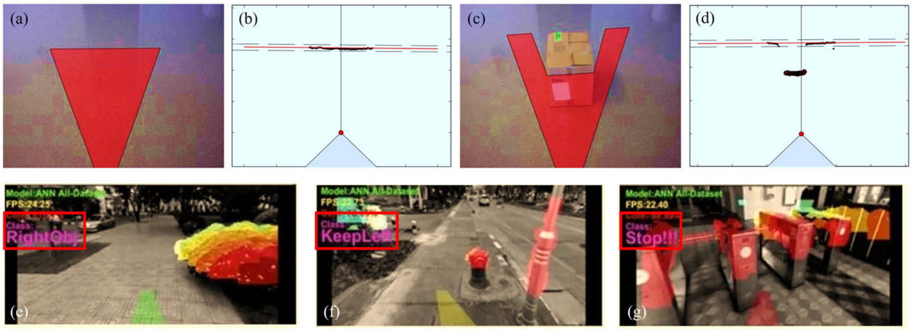

Fig. 5 (a) Whisker logic, as proposed by Pallejà et al. [34], operates here. The LiDAR sensor discerns the flat floor, while the red conical shape signifies the whisker. (b) The graph depicts the flat floor as a straight, horizontal line. (c) A cubic obstacle enters the LiDAR sensor's range, indicating contact with the whisker. (d) The graph illustrates the newly detected obstacle, where the sides are identified as the flat floor, while the central portion displays an indentation denoting an obstacle. (e) Outputs generated via 3D instance segmentation, a method proposed by Vorapatratorn et al. [46], present object detection and navigation instructions. An obstacle is detected on the right, prompting the user to continue straight. (f) Detection of obstacles in the middle and on the right side leads to an instruction for the user to veer left. (g) Detection of an obstacle directly ahead necessitates a stop instruction due to the absence of a feasible route

3D LiDAR

Navigating crowded environments poses a formidable challenge for individuals with visual impairments. In a bid to provide assistance, Kayukawa et al. [22] have devised a mobile system that anticipates pedestrians' trajectories and orchestrates collision-free routes through congested spaces. Their innovation incorporates two RGB depth cameras, a 3D LiDAR, and an Inertial Measurement Unit (IMU) sensor atop a conventional wheeled suitcase, as depicted in Fig. 3(a). Pedestrian detection is accomplished through CNN-based models analyzing video feeds from both cameras. Position and distance calculations rely on depth information from the cameras and 3D point cloud data processed by Simultaneous Localization and Mapping (SLAM) techniques on the LiDAR data. Future pedestrian locations are predicted by continuous monitoring, with the IMU sensor tracking the user's walking pace. The system operates in two modes: on-path and off-path, as illustrated in Fig. 3(c). In on-path mode, walking speed adjusts in response to pedestrians ahead, while off-path mode dynamically computes new routes based on pedestrian positions. Navigational cues are communicated through text-to-speech alerts or haptic feedback, aided by a servo motor and level attached to the suitcase's handle. The servo motor assists off-path mode by guiding the user toward collision-free directions, as seen in Fig. 3(b). The system garnered positive feedback from 14 visually impaired participants, achieving average System Usability Scale (SUS) scores of 76.1 and 75 for audio warnings and haptic feedback, respectively.

While outdoor navigation proves demanding for the visually impaired, independent indoor navigation presents even greater complexity, particularly in identifying furniture like beds, chairs, and tables (Fig. 4 gives an overview of these systems). To address this challenge, Liu et al. [28] have crafted a LiDAR-driven indoor navigation system that not only guides but also identifies and locates furniture within rooms. Employing the Intel RealSense LiDAR camera

L515 (similar to LiDAR system shown in Fig. 5), this system captures high-definition video and 3D point cloud data, complemented by Bluetooth earphones for audio instructions and navigation guidance. User initialization involves issuing a "start scanning" audio command, triggering the capture of surroundings' 3D point cloud and video. SLAM establishes the user's position within the room, and a 3D instance segmentation model processes the point cloud. Combining RGB images from video with 3D point cloud data facilitates object identification, recognizing clusters as distinct objects such as desks, tables, and chairs. This system excels in obstacle avoidance and object identification, guiding users along discerned pathways. Object-specific audio commands allow users to locate items like beds or chairs, enhancing accessibility. Evaluation revealed a ScanNet v2 dataset mean Average Precision (mAP) of 42.4% for their instance segmentation model, while real-life tests with 7 participants garnered an average user rating of 4.57 out of 5.

Setiadi et al. [41] harnessed an RGB camera and 3D LiDAR sensor to detect pedestrian paths and objects for the visually impaired. Pedestrian paths were accentuated through RGB image processing involving color filters, morphological operations, and resizing. A neural network with 4 output classes determined the path direction using a dataset of 400 images. For 3D LiDAR point cloud data, distance-based conversion yielded 2D arrays, which a neural network with 8 output classes employed to detect paths and obstacles. Training and evaluating with 400 LiDAR point clouds, they achieved an 89.7% accuracy for path detection and an 87.5% accuracy for object detection.

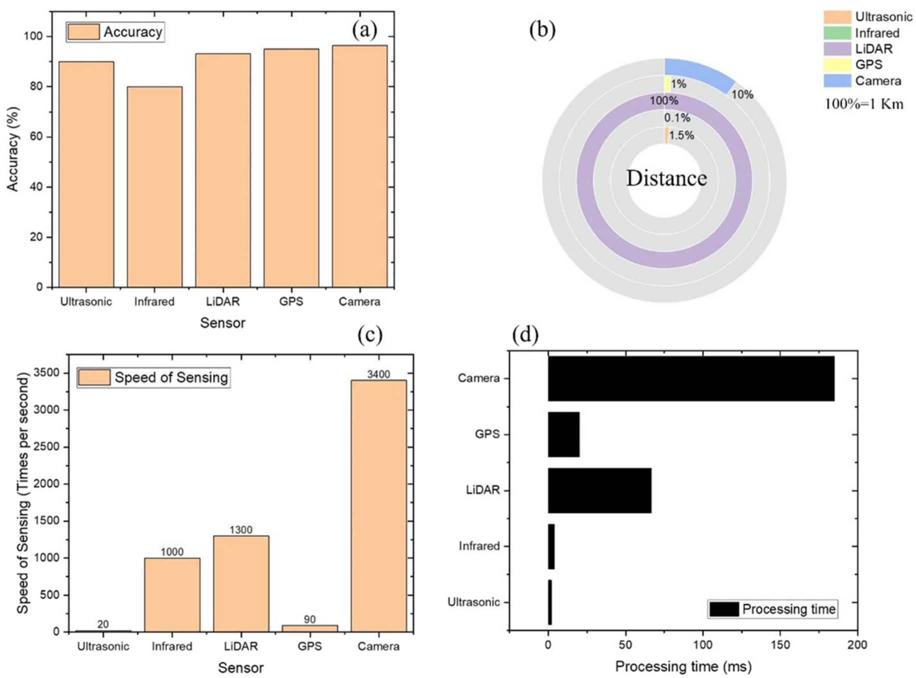

Fig. 6 (a) Comparison of different sensors in blind navigation with respect to the accuracy with which they can predict obstacles. (b) Range comparison of different sensors normalized at 100 %, which is equal to a 1 km distance. (c) Sensing speed for different sensors in times per second. (d) Post-processing time of different sensors

6 Discussions

Figure 6a illustrates the comparison of different sensors in blind navigation concerning their accuracy in predicting obstacles. The results confirm that all five sensors are capable of achieving over 80% accuracy, with the camera, GPS, and LiDAR exhibiting the highest accuracy levels.

Figure 6b depicts the range comparison of different sensors, normalized at 100% equal to a 1 km distance. Notably, LiDAR outperforms all other sensors in terms of range. LiDAR can provide accurate readings even up to 1 km, with an error margin within 2 cm, demonstrating exceptional accuracy. While GPS boasts the maximum coverage range, it remains practical for navigation within only a few meters, and it can only discern movement after a certain distance of travel.

Figure 6c showcases the sensing speed of different sensors, measured in times per second. Remarkably, cameras outperform other sensors in this category, with a maximum scanning speed of 3400 times per second. Although cameras capable of scanning over 1 lac frames exist, they tend to be costly. We have considered typical market values for all sensors, and while more advanced, pricier models may exist, our presented values should remain within a 10% margin of accuracy.

Figure 6d outlines the post-processing time of various sensors, where Ultrasonic and Infrared sensors exhibit superior performance. A smaller post-processing time indicates better performance. Notably, camera-based navigation systems show poorer performance in this parameter.

Upon a thorough review of numerous navigation systems catering to the visually impaired, we observed two primary functions embedded within these systems: obstacle detection and optimal path generation, coupled with the dissemination of pertinent information to the user. A diverse array of sensors contributes to obstacle detection, with ultrasonic sensors and various camera types (RGB, RGB-depth, and stereo) emerging as the predominant choices. Ultrasonic sensors adeptly gauge obstacle distances in the forward direction, whereas RGB cameras excel in obstacle categorization, albeit less effectively for distant objects. For precise obstacle identification and distance calculation, the combined usage of RGB-depth cameras or stereo-configured cameras proves indispensable. While RGB-depth cameras directly offer depth data, stereo image captures necessitate processing for pixel-wise depth derivation. This camera amalgamation, frequently employed in blind navigation systems [22, 37], is somewhat susceptible to fluctuations in ambient lighting conditions, thereby potentially impacting performance, particularly during nighttime or low-light scenarios. This challenge could be surmounted by training image-based obstacle detection algorithms with an extensive dataset encompassing nighttime and low-light samples, fostering robustness in varying illumination conditions. Alternatively, a hybrid approach integrating ultrasonic or laser-based sensors alongside cameras can circumvent this limitation, as these sensors remain unaffected by ambient light.

Another salient obstacle-detection sensor category is LiDAR. Distinct operational principles classify LiDAR sensors into three categories: one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) LiDARs. 1D LiDARs, analogous to ultrasonic sensors, employ laser beams for distance estimation, offering heightened precision over longer distances. Unlike ultrasonic sensors with a 15-degree Field of View (FOV) perpendicular to the sensor, 1D LiDARs possess a negligible FOV, exclusively detecting objects directly ahead.

Building upon this foundation, 2D LiDARs leverage rotating 1D LiDARs to acquire distances within a 360-degree horizontal plane. Their output, a distance array spanning angles from 0 to 360 degrees, embodies a 360-degree Horizontal Field of View (HFOV) and a 0-degree Vertical Field of View (VFOV). 3D LiDARs, with a 360-degree HFOV and non-zero VFOV, extend their laser beams vertically, introducing a dimension to 2D LiDARs. A 3D point cloud, comprising distance, horizontal, and vertical angle values, characterizes the output of 3D LiDARs. Each LiDAR category bears distinctive utilities, from 1D LiDARs estimating front distances to 2D LiDARs constructing top-down room views and 3D LiDARs generating comprehensive 3D surroundings maps for collision-free path planning. Cost-effectively, 1D LiDARs offer affordability, while 3D LiDARs present higher pricing.

Navigation systems for the visually impaired exhibit dual operational modes: some furnish users solely with obstacle distance and direction, omitting navigation directives [9], whereas others facilitate complex guidance by calculating collision-free routes to specified destinations. Notably, Simultaneous Localization and Mapping (SLAM) emerges as a favored algorithm for navigation directives, complemented by alternative methods such as the ROS navigation stack, Artificial Neural Networks (ANN), Augmented Reality (AR), A* algorithm, 3D instance segmentation, and GPS sensors. Inclusive of these systems are Inertial Measurement Unit (IMU) sensors, instrumental in tracking user walking speed and orientation. Such data anticipates future user positions, optimizing path selection. For outdoor navigation, the integration of stereo cameras and 3D LiDARs is crucial for environment comprehension and route suggestions. Conversely, indoor navigation may capitalize on pre-existing room maps, leveraging orientation and user location to aid in reaching specific furniture pieces. Notably, providing obstacle detection and path generation is vital, but the mode of information transmission assumes equal significance. Traditional visual interfaces like screens are ill-suited, rendering audio alerts and haptic feedback the favored mediums for interaction among visually impaired users. This manuscript's surveyed navigation systems predominantly employ audio alerts and commands for user engagement. Notably, innovative approaches have surfaced, including servo-operated levers indicating user movement directions [22] and guiding robots augmenting user mobility [21].

Table 1 presents a comprehensive comparison of navigation systems employing LiDAR sensors in conjunction with other sensor modalities. Notably, certain systems focus exclusively on obstacle detection [5, 9, 12, 23], while others encompass full-fledged navigation capabilities [10, 21, 22, 28]. Among these, the prevalent utilization of 2D and 3D LiDARs stands out in systems offering navigation assistance. RGB cameras and RGB depth cameras emerge as the primary visual sensors in the majority of surveyed literature, accompanied closely by ultrasonic sensors and IMUs. Vision-based obstacle detection predominantly leans towards the YOLO algorithm. For navigation functionality, SLAM emerges as the dominant method of choice. A noteworthy trend in the literature is the widespread adoption of audio alerts and haptic feedback to convey critical information. Notable success stories include Chai et al. [9] and Liu et al. [28], reporting high success rates in participant trials. Additionally, Khaitan et al. [23] achieved remarkable object classification accuracy, adding to the diverse achievements showcased within the surveyed studies.

Table 2 offers a comprehensive comparison of literature devoid of LiDAR sensors. With the exception of Mostofa et al. [32], all the referenced studies possess the capability to compute collision-free pathways for guiding users to their destinations. Predominantly, ultrasonic sensors, RGB cameras, stereo cameras, and IMUs constitute the sensor suite. Notably, YOLO

Table 1 Comparison of navigation systems that use LiDAR sensors

Method	LiDAR used	Sensors	Technique for obstacle detection	Navigation technique	Feedback
Chai et al. [9]	1D	-	IR beam	-	Haptic
Kayukawa et al. [22]	3D	RGB-D, IMU	CNN	SLAM	Text-to- speech, Haptic
Liu et al. [28]	3D	RGB-D	3D segmentation	SLAM	Text-to- speech
Bouteraa et al. [5]	1D	Ultrasonic, IMU	Fuzzy logic classifier	-	Audio alerts
Kayukawa et al. [21]	2D	RGB-D	YOLO	ROS navigation stack	Mobile robot
Khaitan et al. [23]	2D	Ultrasonic, RGB	Sound waves, YOLO	-	-
Chitra et al. [12]	1D	RGB	CNN	-	Haptic
Chen et al. [10]	1D	RGB-D, IMU	-	SLAM	Audio alerts

Table 2 Comparison of navigation systems that use ultrasonic, IR and visual sensors

Method	Sensors	Technique for obstacle detection	Navigation technique	Feedback
Kuribayashi et al. [24]	RGB, IR, IMU	CNN IR beam	AR	Haptic
Rahman et al. [37]	Stereo, ultrasonic	Sound wave, YOLO	Depth Information, GPS	Audio alert
Vorapatratorn et al. [46]	Stereo	Depth Information	ANN	Audio alert
Mostofa et al. [32]	Ultrasonic, RGB-D, RGB	R-CNN	-	-
Toro et al. [45]	Stereo, IMU	YOLO	Depth Information, IMU, mapping	Audio alert
Hakim et al. [16]	RGB-D, Ultrasonic	-	Hector SLAM, A* algorithm	-
Zhang et al. [50]	RGB-D, IMU	-	Depth Information	6-DoF Robot

Table 3 Comparison of different sensors used for blind navigation

Sensor	Accuracy %	Distance	Speed of Sensing	Post Processing time
Ultrasonic	89.9 [37]	15.25 m [48]	20 times/sec [26]	2 ms [8]
Infrared	80 [19]	0.5 m [25]	1000 times/sec [35]	4 ms [30]
LiDAR	93.1 [9]	1000 m [27]	1300 times/sec [1]	66.7 ms [51]
GPS	95 [3]	10 m [38]	90 times/sec [2]	20 ms [31]
Visual sensors (cameras)	96.45 [46]	100 m [18]	3400 times/sec [6]	185 ms [40]

[37, 45] and CNN [24, 32] emerged as the favored techniques for obstacle detection, synergizing seamlessly with the distance estimation prowess of ultrasonic sensors. Navigation functionality spans a diverse array of algorithms, encompassing A* algorithm, SLAM, 3D instance segmentation, depth analysis from RGB cameras, and SLAM. Interaction with users primarily relies on auditory alerts. Notably, Vorapatratorn et al. [46] showcased the highest object classification accuracy, further enriching the gamut of achievements documented within the surveyed literature. Tang et al. [43] have posed a significant query in their research: Do LiDAR sensors offer accurate data under adverse weather conditions? To examine the efficacy of LiDARs in pedestrian and obstacle speed detection, they conducted tests using an autonomous bus within a parking area. The outcomes of their experiment demonstrated that, on days characterized by rain or haze, the precision of LiDAR-derived information diminished in comparison to clear days. Furthermore, they noted a compromised ability of the sensor to estimate obstacle speeds. To address this, they devised three failure models to assess the performance of LiDAR sensors, facilitating calibration for improved outcomes during inclement weather.

The Table 3 provides a comprehensive comparison of various sensors used for blind navigation based on four key parameters: Accuracy (%), Distance, Speed of Sensing, and post-processing time. Each sensor's performance is evaluated in terms of these metrics, allowing for an assessment of their suitability for assisting visually impaired individuals. Ultrasonic Sensor with an accuracy of 89.9%, ultrasonic sensors provide reliable obstacle detection. They can detect obstacles within a range of 15.25 meters. The speed of sensing is 20 times per second, and the post-processing time is remarkably low at 2 milliseconds. Infrared sensors achieve an accuracy of 80% and can sense obstacles within a shorter range of 0.5 meters. They operate at a high sensing speed of 1000 times per second and have a post-processing time of 4 milliseconds. LiDAR sensors offer excellent accuracy at 93.1 % and are capable of detecting obstacles over a substantial distance of 1000 meters. The speed of sensing is impressively rapid at 1300 times per second, and the post-processing time is 66.7 milliseconds. GPS (Global Positioning System) sensors demonstrate high accuracy, reaching 95 %. They have a limited obstacle detection range of 10 meters. The sensing speed is 90 times per second, and the post-processing time is 20 milliseconds. Visual Sensors (Cameras), represented by cameras, exhibit the highest accuracy of 96.45 %. They can detect obstacles within 100 meters. Cameras boast a rapid sensing speed of 3400 times per second, but their post-processing time is comparatively longer at 185 milliseconds. By comparing these sensors across these parameters, it's evident that cameras provide the highest accuracy and sensing speed, while LiDAR excels in detecting obstacles over extended distances. GPS offers good accuracy for outdoor navigation, while ultrasonic and infrared sensors are particularly efficient in terms of post-processing time. This information aids in understanding the strengths and limitations of each sensor, allowing for informed decisions when designing navigation systems for the visually impaired.

A key limitation of using a LiDAR sensor arises when its laser beam encounters a reflective surface, causing deflection that prevents its return to the sensor and thus yields no distance reading. This challenge can be mitigated by incorporating a backup sensor, whose readings serve as an alternative when the LiDAR sensor fails to produce an output.

7 Conclusion

In this manuscript, we undertake a thorough examination of recent technological advancements aimed at aiding individuals with visual impairments in navigation. Historically, navigation has posed significant challenges in the realm of computer science; however, the amalgamation of contemporary sensor technology and computer algorithms has ushered in new possibilities. Encompassing an array of sensors-ultrasonic, infrared, vision, and laser-we elucidate techniques for localization, mapping, obstacle detection, and identification tailored for both indoor and outdoor navigation. Additionally, we delve into methods facilitating interaction between machines and humans.

Our work comprises a systematic review poised to guide future researchers in selecting appropriate sensors aligned with specific applications. While we provide a succinct overview of sensor utilization for general obstacle avoidance and navigation, our primary focus centers on their role in facilitating navigation for the visually impaired. Notably, ultrasonic and LiDAR sensors exhibit promising outcomes; nevertheless, depth-sensing cameras, particularly RGB cameras, emerge as the prevailing choice within navigation systems. These depth-sensing cameras frequently extract both depth and non-depth information, with depth aiding obstacle distance estimation and non-depth information enabling object identification. Leveraging advancements in machine and deep learning, enhanced CNN architectures are increasingly deployed for precise obstacle classification. Furthermore, we underscore recent studies employing localization, mapping, and optimal path planning techniques.

Concurrently, we provide insights into diverse feedback methods designed to apprise the visually impaired of obstacle positions, orientations, and distances. Techniques such as text-to-speech, audio feedback, and haptic cues play pivotal roles in this context.

Our study is dedicated to contributing a comprehensive review, intended to empower future researchers embarking on navigation system development endeavors. By the review's culmination, readers should gain a nuanced understanding of localization, distance mapping, obstacle evasion strategies, and human-machine interaction techniques, thus equipping them to judiciously select the most suitable sensor for their specific application.

Author Contributions Conceptualization was done by I. Patel (IP), M. Kulkarni (MK), and N. Mehendale (NM). All the literature reading and data gathering were performed by IP. The formal analysis was performed by IP. Manuscript writing- original draft preparation was done by IP. Review and editing was done by MK and NM. Visualization work was carried out by IP, MK, and NM.

Funding This research is funded by Somaiya Vidyavihar University (SVU), research project 21-22.

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Conflicts of interest Authors I. Patel, M. Kulkarni, and N. Mehendale declare that there has been no conflict of interest.

Ethics approval All authors consciously assure that the manuscript fulfills the following statements: 1) This material is the authors own original work, which has not been previously published elsewhere. 2) The paper is not currently being considered for publication elsewhere. 3) The paper reflects the authors own research and analysis in a truthful and complete manner. 4) The paper properly credits the meaningful contributions of co-authors and co-researchers. 5) The results are appropriately placed in the context of prior and existing research.

Consent to participate This article does not contain any studies with animals or humans performed by any of the authors. Informed consent was not required as there were no human participants. All the necessary permissions were obtained from Institute Ethical committee and concerned authorities.

References

1. Ahsbahs T, Badger M, Karagali I, Larsén XG (2017) Validation of sentinel-1a sar coastal wind speeds against scanning lidar. Remote Sens 9(6):552
2. Akkamis M, Keskin M, Sekerli YE (2021) Comparative appraisal of three low-cost gps speed sensors with different data update frequencies. AgriEngineering 3(2):423–437
3. Bastida Castillo A, Gómez Carmona CD, De la Cruz Sánchez E, Pino Ortega J (2018) Accuracy, intra- and inter-unit reliability, and comparison between gps and uwb-based position-tracking systems used for time-motion analyses in soccer. Eur J Sport Sci 18(4):450–457
4. Bourne RR, Adelson J, Flaxman S, Briant P, Bottone M, Vos T, Naidoo K, Braithwaite T, Cicinelli M, Jonas J et al (2020) Global prevalence of blindness and distance and near vision impairment in 2020: progress towards the vision 2020 targets and what the future holds. Investig Ophthalmol Vis Sci 61(7):2317–2317
5. Bouteraa Y (2021) Design and development of a wearable assistive device integrating a fuzzy decision support system for blind and visually impaired people. Micromachines 12(9):1082
6. Bylow E, Sturm J, Kerl C, Kahl F, Cremers D (2013) Real-time camera tracking and 3d reconstruction using signed distance functions. In: Robotics: Science and Systems, vol 2, p. 2
7. Cao Z, Liu J, Zhou W, Jiao X, Yang D (2021) Lidar-based object detection failure tolerated autonomous driving planning system. In: 2021 IEEE intelligent vehicles symposium (IV), pp 122–128. IEEE
8. Carullo A, Parvis M et al (2001) An ultrasonic sensor for distance measurement in automotive applications. IEEE Sens J 1(2):143
9. Chai A, Lau BT (2020) Exploring the outdoor performance of a lidar-based ground plane checking system for the visually impaired. EAI Endorsed Trans Pervasive Health Technol 6(23):e2–e2
10. Chen CH, Wang CC, Lin SF (2020) 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. IEEE
11. Chen X, Ma H, Wan J, Li B, Xia T (2017) Multi-view 3d object detection network for autonomous driving. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1907–1915
12. Chitra P, Balamurugan V, Sumathi M, Mathan N, Srilatha K, Narmadha R (2021) Voice navigation based guiding device for visually impaired people. In: 2021 International conference on artificial intelligence and smart systems (ICAIS), pp 911–915. IEEE
13. Chokron S, Kovarski K, Dutton GN (2021) Cortical visual impairments and learning disabilities. Frontiers in Human Neuroscience 15:713,316
14. Frizziero L, Liverani A, Donnici G, Papaleo P, Leon-Cardenas C (2021) Smart cane developed with dfss, qfd, and sde for the visually impaired. Inventions 6(3):58
15. Gbenga DE, Shani AI, Adekunle AL (2017) Smart walking stick for visually impaired people using ultrasonic sensors and arduino. Int J Eng Technol 9(5):3435–3447
16. Hakim H, Fadhil A (2021) 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: ICMMCS 2020, pp 661–672. Springer
17. Innet S, Ritnoom N (2009) An application of infrared sensors for electronic white stick. In: 2008 International symposium on intelligent signal processing and communications Systems, pp 1–4. IEEE
18. Jain M, Patel W (2023) Review on lidar-based navigation systems for the visually impaired. SN Comput Sci 4(4):323
19. Jang YS, Kim SW (2017) Compensation of the refractive index of air in laser interferometer for distance measurement: a review. Int J Precis Eng Manuf 18:1881–1890
20. Javed Z, Kim GW (2022) Panovild: a challenging panoramic vision, inertial and lidar dataset for simultaneous localization and mapping. J Supercomput 78(6):8247–8267
21. Kayukawa S, Ishihara T, Takagi H, Morishima S, Asakawa C (2020) Blindpilot: a robotic local navigation system that leads blind people to a landmark object. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems, pp 1–9

22. Kayukawa S, Ishihara T, Takagi H, Morishima S, Asakawa C (2020) Guiding blind pedestrians in public spaces by understanding walking behavior of nearby pedestrians. Proc ACM Interact Mob Wearable Ubiquitous Technol 4(3):1–22
23. Khaitan R, Rahman MH (2020) Real-time mobility assistance for the legally blind pp 213–219
24. Kuribayashi M, Kayukawa S, Takagi H, Asakawa C, Morishima S (2021) Linechaser: a smartphone-based navigation system for blind people to stand in lines. In: Proceedings of the 2021 CHI conference on human factors in computing systems, pp 1–13
25. Kuriya R, Tsujimura T, Izumi K (2015) Augmented reality robot navigation using infrared marker. In: 2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN), pp 450–455. IEEE
26. Latha NA, Murthy BR, Kumar KB (2016) Distance sensing with ultrasonic sensor and arduino. Int J Adv Res Innov Ideas Educ 2(5):1–5
27. Lemmens M (2007) Airborne lidar sensors. GIM international 21(2):24–27
28. Liu H, Liu R, Yang K, Zhang J, Peng K, Stiefelhagen R (2021) Hida: Towards holistic indoor understanding for the visually impaired via semantic instance segmentation with a wearable solid-state lidar sensor. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 1780–1790
29. Maryono D, Hendix AWA, Harjunowibowo D et al (2021) Object detection based on clustering for assistive technology. Adv Sustain Syst 1(1):10–16
30. Mastorakis G, Makris D (2014) Fall detection system using kinect's infrared sensor. J Real Time Image Process 9:635–646
31. Meier, L, Tanskanen, P, Fraundorfer, F, Pollefeys, M (2011) Pixhawk: a system for autonomous flight using onboard computer vision. In: 2011 IEEE international conference on robotics and automation, pp 2992–2997. IEEE
32. Mostofa N, Feltner C, Fullin K, Guilbe J, Zehtabian S, Bacanlı SS, Bölöni L, Turgut D (2021) A smart walker for people with both visual and mobility impairment. Sensors 21(10):3488
33. Nada AA, Fakhr MA, Seddik AF (2015) Assistive infrared sensor based smart stick for blind people. In: 2015 science and information conference (SAI), pp 1149–1154. IEEE
34. Pallejà T, Tresanchez M, Teixidó M, Palacin J (2010) Bioinspired electronic white cane implementation based on a lidar, a tri-axial accelerometer and a tactile belt. Sensors 10(12):11322–11339
35. Papagianopoulos I, De Mey G, Kos A, Wiecek B, Chatziathasiou V (2023) Obstacle detection in infrared navigation for blind people and mobile robots. Sensors 23(16). https://doi.org/10.3390/s23167198 https://www.mdpi.com/1424-8220/23/16/7198
36. Premebida C, Carreira J, Batista J, Nunes U (2014) Pedestrian detection combining rgb and dense lidar data. In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 4112–4117. IEEE
37. Rahman MA, Siddika S, Al-Baky MA, Mia MJ (2022) An automated navigation system for blind people. Bull Electr Eng Inform 11(1):201–212
38. Real S, Araujo A (2019) Navigation systems for the blind and visually impaired: past work, challenges, and open problems. Sensors 19(15):3404
39. Rogalski A, Chrzanowski K (2002) Infrared devices and techniques. Opto-electron Rev 10(2):111–136
40. Sasaki N, Tomaru S, Nakamura S (2017) Development of inter-vehicle distance measurement system using camera-equipped portable device. In: 2017 17th International Conference on Control, Automation and Systems (ICCAS), pp 994–997. IEEE
41. Setiadi B, Supriyadi T, Nugroho H, Solihin R (2020) Navigation and object detection for blind persons based on neural network. Int J Appl Technol Res 1(1):56–65
42. Sinha BB, Dhanalakshmi R (2022) Recent advancements and challenges of internet of things in smart agriculture: a survey. Future Gener Comput Syst 126:169–184
43. Tang L, Shi Y, He Q, Sadek AW, Qiao C (2020) Performance test of autonomous vehicle lidar sensors under different weather conditions. Transp Res Rec 2674(1):319–329
44. Ton C, Omar A, Szedenko V, Tran VH, Aftab A, Perla F, Bernstein MJ, Yang Y (2018) Lidar assist spatial sensing for the visually impaired and performance analysis. IEEE Trans Neural Syst Rehabilitation Eng 26(9):1727–1734
45. Toro AAD, Bastidas SEC, Bravo EFC (2020) Methodology to build a wearable system for assisting blind people in purposeful navigation. In: 2020 3rd international conference on information and computer technologies (ICICT), pp 205–212. IEEE
46. Vorapatratorn S, Suchato A, Punyabukkana P (2021) Fast obstacle detection system for the blind using depth image and machine learning. Eng Appl Sci Res 48(5):593–603
47. Xu Q, Zhong Y, Neumann U (2022) Behind the curtain: learning occluded shapes for 3d object detection. Proc AAAI Conf Artif Intell 36:2893–2901

48. Yasir M, Lestari IN, Setiawan C, Effendi MR, Hamidi EAZ et al (2021) Design and implementation of the blind navigation aids using ultrasonic sensor. In: 2021 7th International conference on wireless and telematics (ICWT), pp 1–6. IEEE
49. Young J, Simic M (2015) Lidar and monocular based overhanging obstacle detection. Procedia Comput Sci 60:1423–1432
50. Zhang H, Jin L, Ye C (2021) An rgb-d camera based visual positioning system for assistive navigation by a robotic navigation aid. IEEE/CAA J Autom Sin 8(8):1389–1400
51. Zhao X, Sun P, Xu Z, Min H, Yu H (2020) Fusion of 3d lidar and camera data for object detection in autonomous vehicle applications. IEEE Sens J 20(9):4901–4913
52. Zhou Y, Tuzel O (2018) Voxelnet: end-to-end learning for point cloud based 3d object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4490–4499