2NDVISION: A SMART NAVIGATION SUITS TO ENSURE SAFE MOVEMENT FOR VISUALLY IMPAIRED USING MACHINE LEARNING AND INTERNET OF THINGS (IOT)

Zhongxuan Xu[1], Ziqi Zeng[2], Ang Li[3]

[1]Harvard-Westlake Upper School, 3700 Coldwater Canyon Ave, Los Angeles, CA 91604
[2]Oaks Christian High School, 31749 La Tienda Rd, Westlake Village, CA 91362
[3]Computer Science Department, California State Polytechnic University, Pomona, CA 91768

ABSTRACT

In response to the challenges faced by the blind and visually impaired, exacerbated by the increasing prevalence of screen-related visual impairments, this proposal aims to address the shortcomings of existing solutions. The proposed solution consists of a cost-effective and secure approach, featuring a suit equipped with sensors and motors, complemented by a mobile app. This integrated system enhances the navigational experience for the visually impaired, offering key features such as safe traversal in any environment, map-guided navigation, and timely warnings of high-speed dangers. The suit utilizes time-of-flight sensors and vibration motors to convey crucial information to users, including distance, direction, and warnings, through distinct vibration signals. The mobile app serves as a valuable supplement to the suit, providing features such as traffic sign detection and advanced map navigation [8]. This cohesive blend of innovative hardware and software components establishes a comprehensive solution, addressing the urgent need for effective, affordable, and technologically advanced tools for the blind and visually impaired amidst the era of widespread screen usage.

KEYWORDS

IoT, Machine Learning, Computer Vision, Sensory Substitution

1. INTRODUCTION

The growing epidemic of myopia due to increased screen use is leading to more cases of serious visual impairments, including macular degeneration and retinal detachment [1]. With this "epidemic of blindness" particularly impacting the younger generation, novel solutions must be developed utilizing more up-to-date technological advancements beyond conventional options like guide dogs and walking sticks.

However, the high costs and unreliability of current solutions globally discourage officials and policymakers from making necessary enhancements for the blind and visually impaired. For example, assistive reading technology only helps in cases where words need to be distinguished, and cannot help in navigating an environment nor offer warning in dangerous situations [2].

David C. Wyld et al. (Eds): AIBD, MLSC, ACSTY, NATP, CCCIoT, SVC, SOFE, ITCSS -2024
pp. 343-353, 2024. CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.140427

Wearable object detection devices work well in identifying what is in front of the user, but they cannot give accurate assessments of the distance away someone is from obstacles nor can they offer a quick enough warning in dangerous situations [3]. Both solutions listed also cost a lot to buy, which can be not affordable for some.

To tackle the problems of high costs, inconvenience, and safety risks present in current solutions, our proposed approach seeks to provide a cost-effective method for achieving three interconnected navigation objectives for the visually impaired and blind: traversability, map navigation, and emergency warning detection. With this solution, users can experience unparalleled safety, convenience, and reliability at an affordable cost.

Kyushu University's Haptic Vest employs 60 actuators for tactile navigation, providing directional cues through torso vibrations. However, its sensory scope is narrower than the 2ndVision project, which incorporates LiDAR and auditory signals, offering a more interactive and comprehensive navigation aid for the visually impaired [9]. The Smart Cane based on IoT enhances traditional white canes with location and obstacle detection but is limited by sound feedback, especially in noisy environments. In contrast, the 2ndVision project overcomes these limitations with advanced sensors and a multi-modal feedback system, along with mobile app integration for comprehensive functionality. Gollagi's proposed solution involves smart glasses with a Raspberry Pi and a camera module, utilizing sonar sensors and deep learning for obstacle identification, OCR for text reading, and voice commands for user interaction, providing a versatile and adaptive solution.

Our solution addresses the growing myopia-related visual impairments by introducing a cost-effective approach that combines a wearable suit and a blind-accessible mobile app. The suit integrates a camera and 10 LiDAR-based time-of-flight sensors strategically positioned to detect obstacles in all directions by measuring the returning time of emitted light pulses [4]. These sensors, coupled with vibration motors, convey visual information through distinct patterns for navigation, object alerts, and obstacle warnings. The blind-accessible app, developed using Flutter UI development kit and the Dart language, complements the suit with object detection and text-to-speech features. It utilizes map APIs like Google Maps to send directional information, translating it into coordinated vibrations for navigation [10]. In emergencies, the app triggers intense vibrations and text-to-speech alerts based on feedback from the suit's sensors.

Our solution excels in providing three interconnected navigation objectives for the visually impaired: traversability, map navigation, and emergency warning detection [11]. Its cost-effectiveness arises from leveraging modern technologies and an innovative fusion of hardware and software components. In contrast to prohibitively expensive and limited-functionality existing solutions, our approach ensures unparalleled safety, convenience, and reliability at an affordable cost. Compared to traditional options like guide dogs and walking sticks, our solution surpasses by offering comprehensive navigation assistance, real-time environmental awareness, and immediate responses to potential dangers. Overall, our method stands out as an effective, affordable, and technologically advanced solution, addressing the limitations of existing approaches and prioritizing the holistic needs of the visually impaired.

In the object avoidance experiment, three groups of blindfolded testers navigated through a maze: one wearing a suit with vibration motors, one using a walking stick, and a control group with no assistance. The suit demonstrated superior performance in avoiding bumps, with almost 0 collisions compared to canes and the control group. While canes were effective in avoiding bumps, the vest outperformed in overall maze completion time, requiring no user action for object detection. The vest testers spent on average 16.76% more time than cane users and 35.79% more time than the control group.

In the danger warning experiment, the suit with object detection was tested against high-speed balls thrown by a tennis ball machine. The average time gap between motor vibration and ball impact was 3.22 seconds, providing users with over 3 seconds to react to emergency warnings. Even in extreme scenarios with a 2.5-second gap, users still had over 2 seconds to avoid danger, highlighting the effectiveness of the object detection system.

2. Challenges

In order to build the project, a few challenges have been identified as follows.

2.1. Implementing Research Proposal

There are also risks while implementing each step of our goals. We've identified two possible challenges we may face while implementing our research proposal. Firstly, time-of-flight sensors did not work as effectively as we hope because their detection view is very limited and may not cover a wide enough range for the user. The sensor only sends out pulses of rays, which is problematic since the sensor may end up not detecting objects that are to the side of the ray. A possible solution to this problem is to increase the number of time-of-flight sensors and to position them at angles where they can detect a broader range of objects.

2.2. Object Detection

Object detection is also a challenge because it was the most technically challenging part of our suit. We had a bad training model and not enough training data so the outcome of our object detection is an ineffective system that cannot detect certain traffic signs properly. This deficiency would have catastrophic results if this were to rely on actual streets. However, if we were to switch to an image dataset that contains traffic signs specifically, our model would possibly be better at recognizing the things that we want it to recognize. Additionally, we could also adjust the learning rate of the deep learning algorithm to be slower around the minimum values to not surpass the solution by a large margin, which also improved the accuracy of identifications.

2.3. Creating Accessible App User Interfaces

Creating accessible app user interfaces is a challenge due to our lack of perspective. Without any testing from actual blind or visually impaired users, we can only hypothesize about the types of changes we have to make for the app to be accessible. App navigation needs to be integrated into pre-existing accessibility software and text-to-speech instructions should be given to users frequently to not confuse [15]. As developers who haven't lived a life of blindness, it is difficult to develop the app from the perspective of a blind person. A direct solution to this lack of perspective is to let blind users test our app and give feedback, and we would then adjust our app based on the feedback given.

3. Solution

Our solution consists of a wearable suit and a blind-accessible mobile app. The suit, designed to hold the phone in the front pocket, integrates a camera and 10 time-of-flight sensors using LiDAR technology (Figure 1a & 1b). The suit will have 10 time-of-flight sensors placed on every side of the user that detect distances through Light Detection and Ranging (LiDAR) technology that times the returning time of emitted light pulses [13]. Each sensor sends its distance reading to a paired vibration motor, which vibrates with increasing intensity as an object nears the sensor, giving the user a clear understanding of when to avoid objects (Figure 1c).

The blind-accessible app, developed with Flutter and Dart, houses both emergency object detection and map navigation features. The map navigation utilizes the Google Maps API to send directional information to the suit, translating instructions into coordinated vibrations for navigation (Figure 2a & 2b). This allows the user to designate destinations to reach when traveling in a public environment.

To address emergency scenarios, such as approaching speeding vehicles, changes in traffic light, sink holes or bumps in the road, the app has an object detection feature that notifies the user about dangers via TTS and violent vibrations (Figure 2c). When the phone is placed in the front pocket of the suit and the object detection feature is active, this rapid alert system would be detecting and warn the users in real time.

Effective communication is ensured through unique motor vibration patterns for each suit function: map navigation, the object detected rapid object alert, and obstacle proximity warning. Vibration intensities scale with distance, enabling users to accurately gauge their proximity to obstacles and navigate safely.

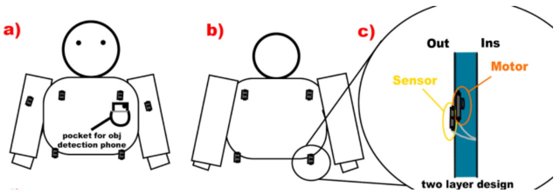

Figure 1. Diagram of the 2ndVision Suit from (a) front view, (b) back view, (c) cross-sectional view

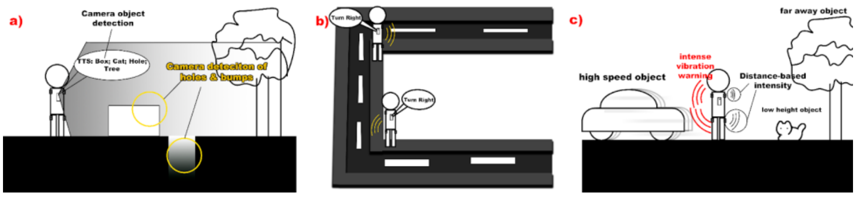

Figure 2. Overview of the functionality of 2ndVision suit (a) object and danger detection, (b) map navigation, and (c) high-speed warning and distance detection

The purpose of the suit component is to enhance the navigation capabilities of visually impaired individuals. The suit's sensors use light detection and ranging (LiDAR) technology to measure the returning time of emitted light pulses for accurate distance detection (Figure 3). Vibration motors paired convey the information from these sensors through Sensory Substitution, changing one's visual modality into touch-based stimuli. Both sensors and motors are connected to an Arduino microcontroller and their behavior is coded in C.

Not applicable but provide diagram of LiDAR:

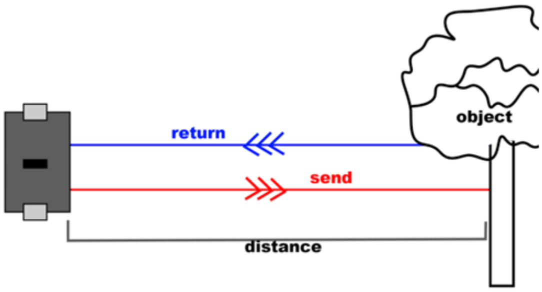

Figure 3. Depiction of how lidar sensor detects distance

```c
if( f_dist > (MAX_FRONT_DIST * 0.75) && f_dist <= MAX_FRONT_DIST * 1){

  Serial.println("Playing haptic_front (75% - 100% of max distance magnitude)");
  play_front(47);
}
else if(f_dist > MAX_FRONT_DIST * 0.5 && f_dist <= MAX_FRONT_DIST * 0.75 ){
  Serial.println("Playing haptic_front (50% - 75% of max distance magnitude)");
  play_front(48);
}
else if (f_dist > MAX_FRONT_DIST * 0.25 && f_dist <= MAX_FRONT_DIST * 0.5){
  Serial.println("Playing haptic_front (25% - 50% of max distance magnitude)");
  play_front(49);
}
else if (f_dist < MAX_FRONT_DIST * 0.25 ){
  Serial.println("Playing haptic_front (0% - 25% of max distance magnitude)");
  play_front(50);
}
```

Figure 4. Screenshot of code in C for suit vibration, numbers in play_front method represent specific waveform intensities

The code displayed controls the front motor's vibration frequency based on the reading from its paired sensor (Figure 4). The front sensor's reading is updated every 100 milliseconds (not displayed) and the if-ladder below categorizes the reading into different ranges of its max detection distance (10 meters) and tells the vibration motor to play waveforms of specific vibration intensities based on which range the distance reading is in. Specifically, for each 25% decrease in the distance from the max detection distance, the vibration motor switches to an increasingly intense vibration.

The blind-accessible app is developed with Flutter, a UI software development kit, and Dart, a programming language [14]. The app houses the object detection feature and a Google ML library's text-to-speech (TTS) functionalities to inform the visually impaired to warn the user of possibly dangerous objects nearby and send signals to the suit via the Bluetooth component to vibrate violently as another layer of alert to the user. Visual indicators of detected objects are shown for the sake of debugging (Figure 5).

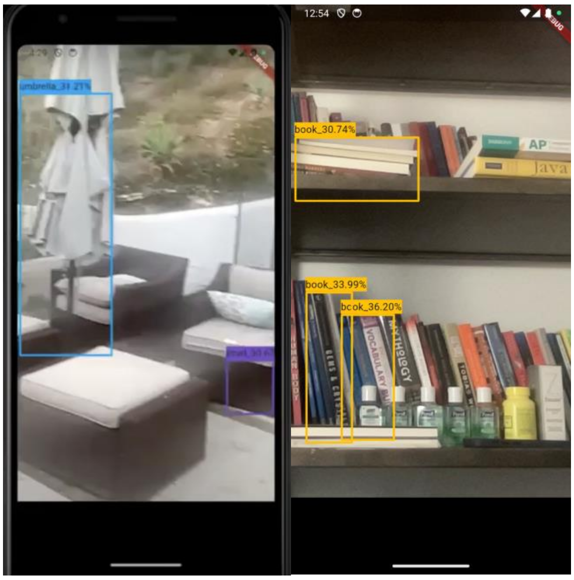

Figure 5. Object detection of exterior items and interior items

```
List<ResultObjectDetection?> objDetect = await _objectModel!
    .getImagePredictionFromBytesList(
        cameraImage.planes.map((e) => e.bytes).toList(),
        cameraImage.width,
        cameraImage.height,
        minimumScore: 0.3,
        IOUThershold: 0.3);

print("data outputted $objDetect");
widget.resultsCallback(objDetect);
```

Figure 6. Screenshot of code in Dart of object prediction where the results are a list of ResultObjectDetection objects called objDetect

The model we created is being used to predict images from the information from the phone camera (Figure 6). Phone camera height and width are specified and we signal a positive detection if objects have an Intersection over the Union threshold (IOU Threshold) of more than 0.3. Resulting predictions, objDetect, is then sent to the resultsCallback method to be formed into the boxes around identified objects seen in Figure 5.

Integrating the Google Maps API, the app sends directional information to the wearable suit via a Bluetooth component, translating instructions into coordinated vibrations for navigation. This streamlined integration allows users to designate destinations, enhancing their ability to traverse public environments effectively.

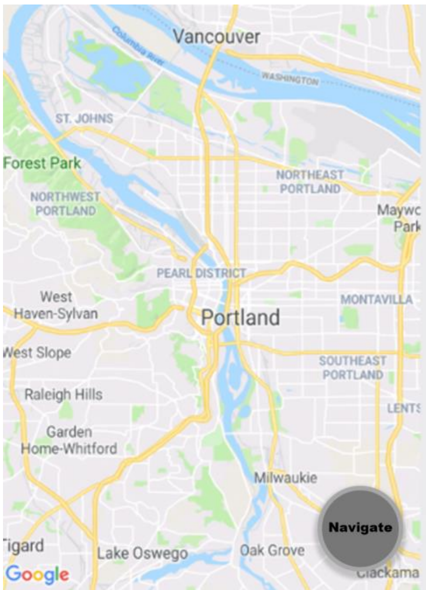

Figure 7. Map navigation displayed for debugging purposes. In practice, directions will be sent to the suit via BlueTooth

```
  body: GoogleMap(
    onMapCreated: _onMapCreated,
    initialCameraPosition: CameraPosition(
      target: LatLng(coordX, coordY),
      zoom: 12.0,
    ), // CameraPosition
  ), // GoogleMap
  floatingActionButton: FloatingActionButton(
    onPressed: () {
     bluetooth(coordX, coordY);
    },
    child: const Icon(Icons.directions),
  ), // FloatingActionButton
); // Scaffold
```

Figure 8. Screenshot of code in Dart for map navigation. This segment is within the UI scaffolding code for the app

The body of the page is the Google Maps display where the map is displayed for debugging purposes. When directions of where to go are entered, they are converted into coordinates and users can then tap on a blind-accessible button placed to the bottom right (floatingActionButton) corner to call on the navigation. At that point, the directions are sent via the Bluetooth component of the suit to the motors using the Bluetooth () method. Specifically, the Google Maps API directions are translated into vibrations in motors on different sides of the suit depending on the direction the API wants the user to head, allowing the user to follow visual instructions even though they are blind [12].

4. EXPERIMENT

4.1. Experiment 1

To prove the effectiveness of our solution, we devised an experiment to evaluate the effectiveness of our suit's object-avoiding capabilities.

We used a validation experiment for the object avoidance of the suit by constructing a simple maze made of cardboard boxes of varying heights and letting blindfolded testers walk through the maze. Three different groups of blindfolded testers participate in this experiment: a group that wears the suit to navigate through the maze, a group that uses a walking stick to navigate through the maze, and a control group that has no assistance. We measured the amount of times that each group bumped into the cardboard walls of the terrain and the time each participant took to complete the maze. We then graphed the bump data for each group on a distribution graph and the average time and the time each participant spent traversing through the maze on a separate participant/trial # vs time in seconds graph. This way, we can compare the suit's improvements to both walking stick traversal and purely blind traversal.

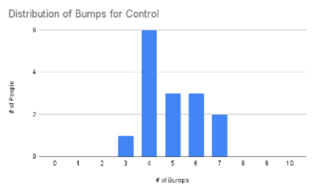

Figure 9a. Distribution of Bumps of Participants for Control group

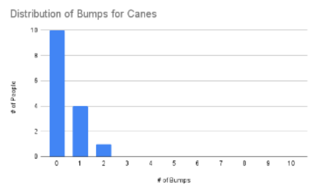

Figure 9b. Distribution of Bumps of Participants for the Canes group

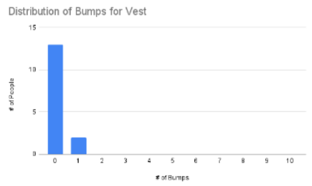

Figure 9c. Distribution of Bumps of Participants for Vest Group

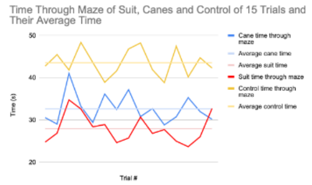

Figure 9d. Time through a maze of suit, maze and control over 15 trials and each group's average time across trials. Each trial's data is the timing for a single participant.

The bump distribution showed a clear advantage for vests thanks to the immediate signaling of the vibration motors. With the control group having on average 4 bumps, the vest's performance of almost 0 bumps shows a significant relative increase in a blind person's ability to avoid obstacles. Canes also performed solidly in avoiding bumping into walls, which is justified by the canes ability to also send immediate feedback to users when the cane bumps into something. However, the advantage of the canes wanes when we factor in the time in which a tester can complete a maze. Although canes are great at searching for objects in their near vicinity, it is too slow when compared to the vest, which requires no action on the side of the user to detect objects. This distinction allows vest testers to walk through the maze without constantly tapping for obstacles, greatly decreasing the time they spend on the maze. Our data further justifies our conclusion that the vest performs better at object avoidance than existing solutions (e.g. canes): the average time spent on the maze for vest testers is on average 16.76% higher than canes users, and is 35.79% higher than the control group.

4.2. Experiment 2

Not only do we have to guarantee the effectiveness of the suit, the most salient feature of our app should also be tested. Namely, we want to make sure that our emergency notification works in a real-world scenario, giving the user enough time to react.

We used a verification experiment for the danger warning experiment. There will be no testers involved with this experiment. We will place the suit with the object detection feature on the app open on a dummy doll and use a tennis ball machine to throw high-speed balls at the suit. We will be measuring the time at which the high-intensity warning vibrations occur versus when the speeding ball hits the suit. This allows us to measure the amount of time the user has to avoid danger so we can properly verify the effectiveness of our object detection.

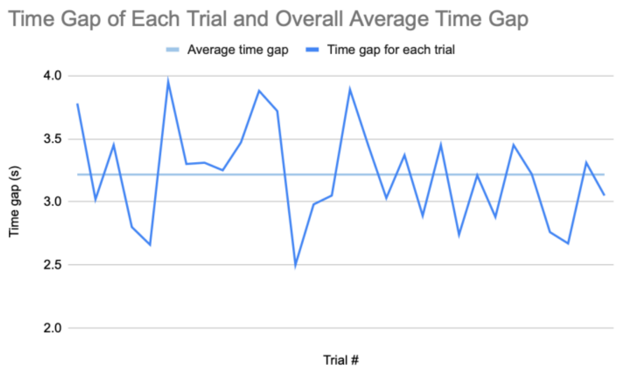

Figure 10. Graph of time gap of 30 trials and overall average time gap

The average time gap between the motor vibration and the impact of the tennis ball is around 3.22 seconds, which gives the user a short yet not impossible period to react to an emergency warning such as dodging any incoming speeding vehicle or stopping at a red light. Given that human reaction time is faster with touch stimuli such as our vibration motors (around 0.15 seconds), the user has a little over 3 full seconds to make a response after sensing the vibration. Even in the most extreme scenarios in our testing where the time gap is only 2.5 seconds, the user still has over 2 seconds to avoid danger, still a relatively long time in terms of rapidly occurring events.

5. Related Work

Kyushu University's Haptic Vest uses 60 actuators for tactile navigation, providing directional cues through torso vibrations [5]. Although innovative in tactile feedback, it has a narrower sensory scope compared to the 2ndVision project. 2ndVision incorporates a wider range of sensory feedback, including LiDAR and auditory signals. It also enhances user experience through a mobile app that offers dynamic navigation, map integration, and emergency alerts. This extensive approach, integrating advanced technology and user-friendly features, offers a more interactive and comprehensive navigation aid for the visually impaired, surpassing the Haptic Vest's capabilities.

Smart Cane based on IoT Faiz Bin Abdul Ghani, and Asuna Zulfadhli bin Zariman1 enhances traditional white canes by offering location and obstacle detection [6]. However, its reliance on sound for feedback limits its effectiveness, especially in noisy environments and it lacks indoor navigation and above-ground obstacle detection. Conversely, the 2ndVision project surpasses these limitations with a sophisticated array of sensors, including LiDAR, and a multi-modal feedback system combining vibration and auditory signals. This approach, coupled with its mobile app integration, enables advanced functionalities like map navigation and emergency alerts, offering a more comprehensive and adaptable solution for the visually impaired.

The solution proposed by Gollagi in the Indonesian Journal of Electrical Engineering and Computer Science involves the use of smart glasses equipped with a Raspberry Pi single-board computer and a camera module [7]. This system captures environmental images and uses a sonar sensor interfaced with a Node MCU to detect obstacles near the blind user. The data from the sensor and the camera are correlated to identify obstacles. The system employs deep learning models to process the captured images, enabling the identification of various obstacle types. It also includes an Optical Character Recognition (OCR) feature to read text on boards for the user, and it operates with voice commands for user interaction.

6. Conclusions

Our approach is limited by the battery life of our suit. If the suit runs out of battery, all features will cease to function, which can be detrimental if it happens at the wrong time during the day. We plan to either add larger batteries for the suit to run or give plenty of low battery warnings so the user knows to recharge the battery.

Additionally, though our suit's object avoidance is satisfactory, we still strive to increase the time gap of notifications before dangerous events occur. Although 2-3 seconds of response time is relatively long, we have to consider all possibilities of human error as it is very likely that real users will freeze in shock or fear and waste the time gap provided by the notification.

Although we've referred to this project as a "solution" or "tool," this is more than innovation; it's a commitment to the holistic well-being of the visually impaired and blind community. With this suit, we turn darkness into a landscape of tactile awareness, where every touch illuminates a path forward.

References

[1] Haarman, Annechien EG, et al. "The complications of myopia: a review and meta-analysis." Investigative ophthalmology & visual science 61.4 (2020): 49-49.

[2] DEMIRAL, Dilek GONCER. "Emerging Assistive Technologies and Challenges Encountered." Current Studies in Technology, Innovation and Entrepreneurship (2023): 1.

[3] Chakraborty, Tusher, Taslim Arefin Khan, and ABM Alim Al Islam. "Flight: a low-cost reading and writing system for economically less-privileged visually-impaired people exploiting ink-based braille system." Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017.

[4] Wandinger, Ulla. "Introduction to lidar." Lidar: range-resolved optical remote sensing of the atmosphere. New York, NY: Springer New York, 2005. 1-18.

[5] LI, Yan, et al. "A design study for the haptic vest as a navigation system." International Journal of Asia Digital Art and Design Association 17.1 (2013): 10-17.

[6] Ghani, Faiz Abdul, and AsnazulfadhliZariman. "Smart cane based on IoT." International Journal of Education, Science, Technology, and Engineering (IJESTE) 2.1 (2019): 12-18.

[7] Rajendran, P. Selvi, Padmaveni Krishnan, and D. John Aravindhar. "Design and Implementation of Voice Assisted Smart Glasses for Visually Impaired People Using Google Vision API." 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2020.

[8] Joorabchi, Mona Erfani, Ali Mesbah, and Philippe Kruchten. "Real challenges in mobile app development." 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. IEEE, 2013.

[9] Behroozpour, Behnam, et al. "Lidar system architectures and circuits." IEEE Communications Magazine 55.10 (2017): 135-142.

[10] Chow, T. Edwin. "The potential of maps APIs for internet GIS applications." Transactions in GIS 12.2 (2008): 179-191.

[11] Rohs, Michael, et al. "Map navigation with mobile devices: virtual versus physical movement with and without visual context." Proceedings of the 9th international conference on Multimodal interfaces. 2007.
[12] Hu, Shunfu, and Ting Dai. "Online map application development using Google Maps API, SQL database, and ASP .NET." International Journal of Information and Communication Technology Research 3.3 (2013).
[13] Reutebuch, Stephen E., Hans-Erik Andersen, and Robert J. McGaughey. "Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory." Journal of forestry 103.6 (2005): 286-292.
[14] Ferreira, Jennifer, James Noble, and Robert Biddle. "Agile development iterations and UI design." Agile 2007 (AGILE 2007). IEEE, 2007.
[15] Coutrot, Antoine, et al. "Virtual navigation tested on a mobile app is predictive of real-world wayfinding navigation performance." PloS one 14.3 (2019): e0213272.