A Smart Walking Aid for the Visually Impaired: A Case Study of the University of Ghana

Robert A. Sowah, Percy Okae, Deborah A. Martin, Abdul Aziz A. Tetteh, Christopher K. Fiah, and Sedinam B. Dwamena

Department of Computer Engineering, University of Ghana, Legon.Accra, Ghana ***Corresponding author:** rasowah@ug.edu.gh

ABSTRACT

Mobility is one of the biggest challenges of the visually impaired and recent studies focusing on how technology can be used to mitigate the effects of this challenge have resulted in the development of smart canes. Current innovations of smart canes include sensors that detect obstacles as well as voice-controlled navigational capabilities. These innovations do not necessarily inform the user of what the potential obstacle is and also require an internet connection for navigation. This paper describes an offline smart cane, called WalkMaTE, with obstacle detection and classification, as well as voice, controlled navigational capabilities. The system is equipped with a pair of ultrasonic sensors placed strategically to detect branches above and obstacles below. Also, an infrared sensor is connected to a vibrating motor which detects nearby gutters and alerts the user(s). The system takes images of the user's surroundings using a Raspberry Pi camera and determines if there are any potential obstacles in the captured image using TensorFlow (an open-source library). To have a sense of where the user is at any point in time, we use a Global Positioning System receiver module to get the coordinates which are then inputted into the navigational system. The destination of the user is provided using speech through a Bluetooth headset with the help of the CMUSphinx speech library. Routing is done with an open-source navigational system called NAVIT. The uniqueness of WalkMaTE is that it works without the use of the internet which is helpful to the visually impaired living in developing countries where the internet is not readily available. The prototype of the smart cane helps the user to get to unknown places on the University of Ghana campus and aids with the identification of obstacles in the path of the visually impaired.

Keywords: navigation; NAVIT; offline smart cane; raspberry pi; TensorFlow; voice control and speech recognition

Introduction

Visual impairment is the loss of vision of a person or a significant limitation of visual capability. Vision is the most essential part of the human physiology as 83% of the information human beings get from the environment is via sight (Agarwal *et al.*, 2017; Aymaz & Cavdar, 2016; Hashino & Ghurchian, 2010; Jose *et al.*, 2016; Kumar *et al.*, 2014). In Ghana, data gathered suggests that there were over 190,000 blind persons in Ghana in 2016 including students of the University of Ghana. The visually impaired can move about with the help of walking aid tools which have evolved from the use of

humans as guides to the use of specially trained guide dogs. Through this evolution came the white cane that helps the blind to determine what is in front of them by swinging the stick from side to side. All these tools above have certain limitations. For example, the use of a human guide limits the blind from living independently. Though the introduction of white cane addressed the privacy issue, it still has some drawbacks such as its limited ability to detect objects within the immediate vicinity of the person and cannot determine the type of object for possible maneuvering.

Furthermore, the white cane is unable to detect obstacles such as low hanging branches and in some cases gutters. The latest innovation in walking aids is the smart walking aid where sensors such as infrared and sonar sensors are mounted on walking sticks or bands to detect obstacles. However, there is still the disadvantage of not knowing the type of obstacle. For the blind, navigating to unknown places can be particularly tricky as they are unable to see signboards that would give indications of where they are or how to get to their respective destinations. In the University of Ghana setting, for example, one of the problems highlighted was the problem of not knowing how to get to unknown examination halls without the help of companions who might also be busy during that period. Blind people, therefore, face the risk of getting lost and are restricted to travelling to places they are familiar with.

Traditional white canes are only able to detect objects within the vicinity of the holder giving barely enough time to move away from the obstacle. They do not inform the holder of approaching obstacles such as other people and silent cars and could result in serious injuries. Most smart canes available are developed for countries with readily available internet and thus, are ill-equipped for a developing country such as Ghana where internet connectivity and penetration are not ubiquitous.

The specific objectives of this research work were to design and develop an intelligent walking aid that: (1) determines where a person is and directs the person to their desired destination through speech recognition; (2) detects incoming obstacles as well as low hanging tree branches and with the capability of informing the user the nature of the obstacle for possible avoidance; (3) detects uncovered gutters; and (4) ensures system's possible operation without internet usage.

Review of Existing Works on the Design of a Smart Walking Aid

Jose *et al.* (2016) proposed a scheme that provides a modest budget and efficient navigational aid for the blind. This system consists of a simple walking stick equipped with ultrasonic sensors to give information

about the environment such as obstacles, pits, and puddles ahead of the user. This system also has a Global Positioning System (GPS) technology integrated with preprogrammed locations for navigation. The user chooses the destination from a set of locations stored in memory. Also, voice-enabled switching equipment is included. But the system does not provide information on the type of obstacle detected, does not determine the speed of an approaching obstacle, is unable to identify hidden obstacles like downward stairs, and in noisy environments, the user would be unable to hear the voice alerts.

Nada *et al.* (2015) proposed a solution in the form of a walking stick with infrared sensors to detect staircases and a pair of ultrasonic sensors to detect other obstacles in front of the user within the range of 4 meters. Also, a water sensor is used to detect the presence of water. Voice alerts and the vibration motor are activated whenever an obstacle is detected. The system is also equipped with a radio frequency (RF) transmitter and receiver to find the stick if it gets missing. Although speech alerts are given, the type of obstacle is not given.

A lightweight, cheap, user-friendly, fast response and low power consumption, smart stick based on infrared technology has been proposed (Agarwal *et al.*, 2017; Jose *et al.*, 2016; Nada *et al.*, 2015; Sen *et al.*, 2018; Tudor *et al.*, 2015). The developed systems by all these authors have a pair of infrared sensors strategically placed to detect staircases and other obstacles within a 2-meter range. Infrared sensors are not highly immune to noise; hence the noise could interfere with the reading and corresponding response to detected obstacles for possible avoidance.

Gayathri *et al.* (2014) proposed a system that consists of two units, the sensor unit, and the Global Positioning System (GPS) unit. This system alerts visually impaired people over obstacles, pit, and water and outlines a better navigational tool that consists of sensors that give information about the environment.

The system uses an ultrasonic sensor, a pit sensor, a water sensor, a GPS receiver, a level converter, a driver,

a vibrator, a voice synthesizer keypad, a speaker or headphone, a PIC microcontroller, and battery. In this system, the visually impaired person can travel only up to four locations using the stick, and the type of obstacle is not known.

Based on the above review and having identified key weaknesses and strengths, we proposed the design, development and deployment of WalkMaTE, a voicecontrolled navigational system with the capability for object detection and obstacle avoidance to aid the visually impaired. The two (2) key features of this system include:

(*i*) *Voice Controlled Navigational System-* This is intended to map out the University of Ghana campus using an open-source offline library called NAVIT taking into consideration that the internet in Ghana is not ubiquitous and readily available on demand. Thus, the users would indicate their destinations into the Bluetooth headset. That information is sent to the walking aid through Bluetooth, and the information is converted to text. The software on the stick then gets the location of the person, calculates the distance and direction, and communicates it to the person through the Bluetooth headset, having converted the text to speech.

(ii) Obstacle Detection and Classification- This system makes use of two ultrasonic sensors and a camera on the stick which is needed for communication with the Raspberry Pi. The upper ultrasonic sensor detects objects which are above the user (preferably the upper body part of the user) such as branches from a distance and the lower ultrasonic sensor detects objects which are ahead of the user at a particular distance. When an obstacle is detected, the camera takes an image of that object and sends the image to the Raspberry Pi for recognition and classification using an open-source library called TensorFlow from Google retrained with a database of images from the specified use case (TensorFlow, 2018). The resulting information is converted to speech and communicated through voice to the visually impaired through the Bluetooth headset. Also, there is an infrared sensor which is used for gutter detection. In the presence of a gutter, the vibration motor vibrates to alert the user of possible gutters.

System Design and Methodology

Figure 1 shows the block diagram of the integrated system design with different components and their interactions. The Global Positioning System (GPS), the camera, and the ultrasonic sensors act as inputs to the Raspberry Pi 3 microcontroller. The output of the Raspberry Pi microcontroller is the headset connected via Bluetooth to the walking stick. The infrared sensor act as input to the Arduino, and the output is the vibration motor which aids in obstacle avoidance and alerts.

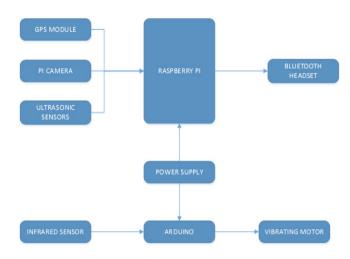


Fig. 1: Block diagram of the integrated system

Obstacle Detection and Classification System

The obstacle detection and classification system are designed by using the ultrasonic sensor, text to speech library (Flite) and TensorFlow. The ultrasonic sensor has a range from 2cm to 400cm and a 15 to 30° measuring angle. Since the ECHO output is 5V and the input pin of the Raspberry Pi GPIO is 3.3V, a voltage divider circuit was used to bring down the voltage to 3.3V. Thus, the voltage divider formula is used.

$$V_{out} = \frac{R_1}{R_1 + R_2} V_{in}$$

The various resistor components for the circuit design were appropriately selected. There were two ultrasonic sensors, one for detecting obstacles like branches and the other for detecting other obstacles like chairs, cars, and so on, hence, the upper and lower ultrasonic sensor. The lower ultrasonic sensor triggered the camera to take a picture. The distance was calculated by the product of the speed of sound and the time it takes for the echo to bounce back divided by two.

$$Distance = \frac{speed \times time}{2}$$

When there was no obstacle detected or the obstacle fell out of range, the ultrasonic continuously detects and senses objects. For the lower ultrasonic sensor, when there was an obstacle, the message "Obstacle Ahead, watch your steps" is read out to the hearing of the person before the type of obstacle was read out and vice versa for the upper one. The Obstacle Classification is handled by TensorFlow, which is an open-source software library offered by Google, for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them (Tensorflow, 2018). The need to use neural networks for this aspect was because of its ability to learn from the training set without necessarily memorizing the training set, making it biased and not applicable to reallife situations. Using Inception-v3, a pre-trained model, trained by validating against ImageNet - an academic benchmark for computer vision, the transfer learning technique was used to achieve the purpose for the specified use case of the University of Ghana. Modern object recognition models have millions of parameters and can take weeks to train fully. Transfer learning is a technique that circumvents or shortcuts a lot of this work by taking a fully trained model for a set of categories like ImageNet and retrains them from the existing weights for new classes (Demirovic et al., 2018; TensorFlow, 2018). Before training, a set of images was needed to teach the network about the new classes to recognize in the specified use case. Over 3000 images of objects recognized as obstacles were captured and classified to be used. The first phase analyzes all the images supplied and calculates bottleneck values for each of them. Once the bottlenecks are complete, the actual training of the top layer of the network begins. Randomly selected

images from a different set are used in the validation after training to prevent the memorization of the dataset and encourage the model to learn and gain the confidence of its prediction over time. Cross-entropy is a loss function which gives a glimpse into how well the learning process is progressing. The objective of the training is to make the loss function as small as possible. As a result of some limitations on the Raspberry Pi, there was the need to optimize the model to run smoothly per the application for the specified use case. The inception model includes a DecodeJpeg operation hence the need for the optimization. The script performs other optimizations that help speed up the model, such as merging explicit batch normalization operations into the convolutional weights to reduce the number of calculations. Hence, by using the optimized graph generated from the optimized tool, the speed was improved for the purpose intended for WalkmaTE. The top five scores with labels were printed out during testing. The topmost label with a probability of 0.5 and above is piped out as speech using Flite text-to-speech through a Bluetooth headset connected to the Raspberry Pi microcontroller. For example, if a chair is recognized, the resulting speech is "Approaching a chair," if the probability is greater than 0.5 but, "Probably approaching a chair," if the probability is less than 0.5. The Pi camera is strategically placed on the walking aid to get a good viewing angle and classify accurately the images it captures if the ultrasonic sensor detects an obstacle based on its calibration. Flite library is embedded into the object recognition and classification to convert the text of classified objects to speech for the blind. Figure 2 shows the flow diagram of the Obstacle Detection and Classification System.

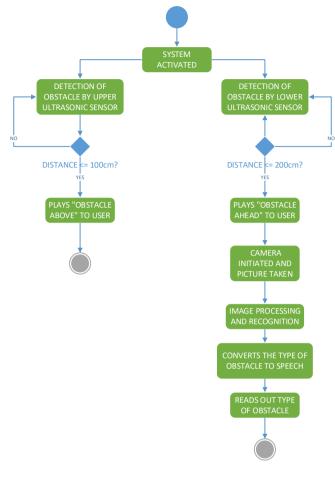


Fig. 2: Flow diagram of Obstacle Detection and Classification System

Voice Control Navigation System

The voice control navigation system involved a turn by turn navigation aid to the visually impaired. It is made up of *Speech-To-Text, Text-To-Speech,* and Maps. CMUSphinx is used for the Speech to Text and to make it more accurate; a grammar model was built. A language model was also required because the pre-trained model was extensive as it contains all English words. A different text-to- speech synthesizer was used for this part, this is known as *eSpeak* (Blasch *et al.,* 2015) (Lakdawala *et al.,* 2018; Prakoso *et al.,* 2016; Shivakumaret *et al.,* 2017). For the navigation system, NAVIT (GitHub, 2018; NAVIT, 2018) were used, and the steps used to achieve the smart cane development goals are as follows: *Setting the map* - Because the University of Ghana campus so happened to be our case study, the map of the University was downloaded from OpenStreetMaps.

Getting the current location of the user - A GPS module receiver was added to the project to get the current location of the user. The mathematical principle used here is trilateration.

Setting the destination - NAVIT does not come equipped with the ability to set the destination through speech, which is an essential part of the project. With the help of the NAVIT team, an interface was written using the dbus to help communicate to NAVIT through a thirdparty script. That was done to enable the customization of NAVIT such that voice control could be added to the work as it was not developed for the visually impaired. The extension helps to set destination using bookmarks which are already stored in NAVIT containing the locations of consideration under the desired use case using a Python script.

Routing - The routing takes advantage of NAVIT's turnby-turn navigation to aid the visually impaired to the desired location. NAVIT makes use of the waypoint technique to connect waypoints along the path from the current location to the desired destination. Waypoints are simply longitudes and latitudes of positions on the surface of the earth. After using the waypoint technique, Dijkstra's algorithm is then used to determine the shortest path from the current location to the destination.

Turn by turn navigation - By making use of a text to speech module called espeak, the system can read out updated turn by turn navigation routes to the user by interfacing with the Bluetooth headset. Figure 3 shows the flow diagram for the Voice Control Navigation System.

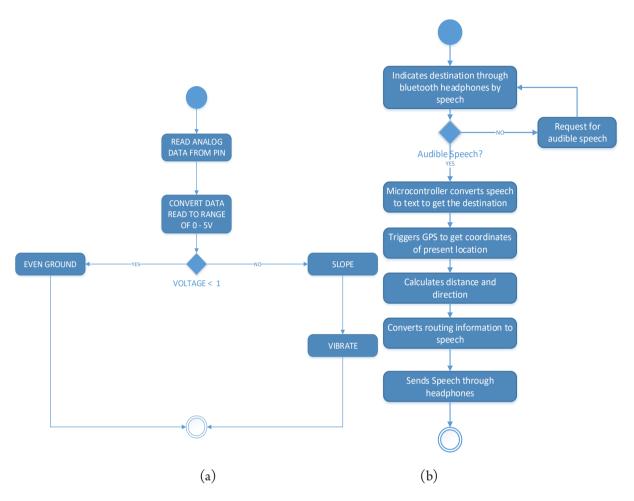


Fig. 3: (a) Flow diagram of Gutter Detection (b) Flow diagram of Voice Control Navigation System

Gutter Detection

The gutter detection involved the Infrared Sensor and the Vibration Motor. The general concept took advantage of the laws of reflection of light. We set the distance of detection of the receiver of the IR to an acceptable distance using the potentiometer. When light bounced off an obstacle within the range, it is assumed to be an even ground. If nothing was sensed, we assume it is just a ditch ahead, and the vibration is sensed by the blind person to take notice of the gutter ahead. The strength of the vibration (PWM) is 80%. Figure 4 shows the flow diagram of the Gutter Detection.

System Hardware Design

Figure 4 shows the front and rear ends of WalkMaTE, and Figure 5 shows the whole circuit design of WalkMaTE.

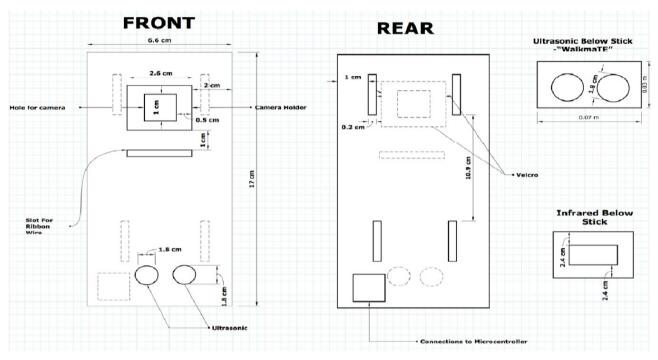


Fig. 4: Front and Rear of WalkMaTE

The circuit design of WalkMaTE in Proteus hardware simulation software.

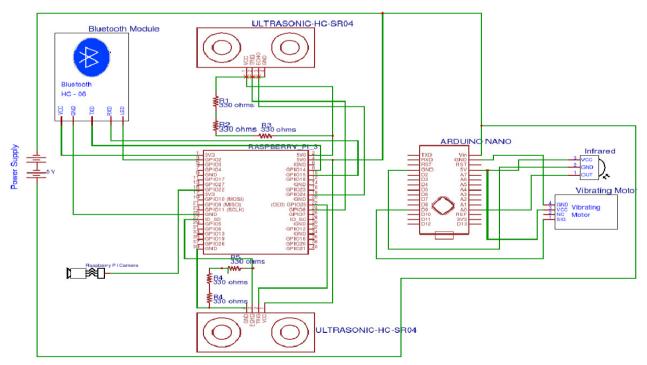


Fig. 5: Circuit diagram developed in Proteus simulation software

System Integration

For the gutter detection, the microcontroller used was the Arduino NANO, and thus there was no need to integrate it with the other parts. It only shared the same power supply with the Raspberry Pi and other components. The Voice-Based Navigation System had the main script that connected all its various parts (i.e., speech to text, text to speech, NAVIT and the GPS). The script for the classification and recognition was written in C++ due to its efficiency and optimization, and that of the ultrasonic sensor was written in Python. Thus, there was a single Python script that connected both scripts for the Obstacle Detection and Classification module. The Obstacle Detection and Classification script is always readily available, but that of the Navigation was only called when the user presses the push-button for navigational information.

Results and Discussion

The results were tested and verified to ensure that the research achieved its stated objectives of providing a smart walking cane for the blind with full navigational capabilities.

Object Detection and Classification

The device was tested in a classroom and room setting. Due to the position of the camera, the system did not always give an accurate instance of what the object was actually in front of the person. It informed the user of the fact that there is an obstacle ahead and based on its classification if the probability was above 0.5 it notified the user of what obstacle was in front of the user. The position of the camera was vital in this aspect of object detection and classification. If the camera is very high, the pictures taken will not correspond to the obstacle in front of the user.

Table 1: Results of object classification and recognition

Actual obstacle	Obstacle detected by system
Monitor	Laptop
Speaker	Speaker
Chair	Chair
Person	Person
Barrel	Dustbin
Bucket	Barrel
Dumb-bells	Dumb-bells

Table 1 shows the results of the Image Classification and Recognition. Due to the picture that was taken by the camera, the obstacle detected might not be the same as the actual obstacle. For instance, the obstacle Barrel was seen as a Dustbin because some of the pictures used to train the classifier have the same features as that of a dustbin. Also, for the Bucket, due to the picture that was taken, there was a barrel in the picture as well as the bucket, and hence the classifier picked the barrel as one with the highest probability.

Table 2: Results of distance covered

No.	Measure Distance (cm)	Sensor Detected Distance (cm)	Error (cm)
1	0	0	0
2	5	4	-1
3	10	8	-2
4	15	14	-1
5	25	23	-2
6	50	49	-1
Average Error			-1.1667

Table 2 shows the difference between the ultrasonic sensor measurement and the actual measurement between the obstacle and the ultrasonic sensor. The average measurement error was 1.1667.

Navigation

Due to the language model and grammar file, the pronunciation of the locations had to be the same as that in the files. Due to this, the system might not respond or pick the first set of words spoken hence a message is repeated for the user to confirm what he or she said. Based on the routing, the user is guided to the destination.

Table 3: Results of speech to text

Location	Number of trials	Breakdown
Engineering	3	'to,' 'engineering.'
NNB	2	'to', 'n', 'n', 'b'
Special needs	3	'to,' 'special,' 'needs.'

Table 3 shows the number of trials the system listens to the user for input and the breakdown of the words for input to the Python script for routing.

Figure 7 shows the navigational system from a source to the destination. After listening to the destination inputted by the user through speech, the system broke the words down and based on its scripts files and bookmarks in NAVIT, the routing part is set, and the description is read out to the user via the Bluetooth Headset.

Overall Hardware Design

Figure 8 shows the overall hardware design of the developed prototype with the various components and where they are placed within the modular hardware design and development of the smart walking cane for the visually impaired and the blind.

Fig. 6: Voice Command for WalkMaTE navigation

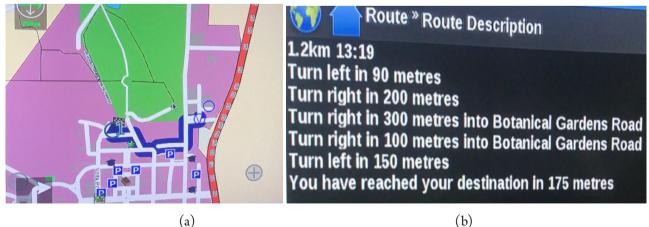


Fig. 7: (a) The path from source to destination on UG Campus (b) Routing Information



Fig. 8: Overall Hardware Design with integrated software system

Conclusion

The paper presents the design of a hardware prototype of smart walking cane that included the combination of a pair of ultrasonic sensors, an infrared sensor, a global positioning system module, vibration motor, Arduino NANO, Raspberry Pi 3, Raspberry Pi camera for the visually impaired. The two primary objectives of this research paper were to increase mobility and enable an offline navigation system. To improve the movement for the visually impaired, a pair of ultrasonic sensors, a Raspberry Pi camera and the infrared sensor was used. The visually impaired can now identify objects and gutters in their path. Implementation of an offline navigation system enables a visually impaired person to move into unfamiliar places without human guidance, thus, making the visually impaired more independent. The hardware and software part of this project has been integrated to meet all the design requirements.

Given that the system is offline, it is clear that all of the challenges cannot be addressed entirely without significant cooperation and an increase in the speed of the image classification. Adjusting the ultrasonic sensor to detect fast-moving objects and adjusting the infrared sensor to detect staircases are recommended. Also, increasing the grammar file to accommodate more pronunciations of locations will help in directing the person to his or her destination as quickly as possible. As the system uses a power bank to supply power, it is recommended that a backup power supply plan on standby to ensure constant power supply. To improve the performance of the navigational system, a Global System for Mobile Communication (GSM) module can be incorporated to improve indoor navigation.

References

- Agarwal, R., Ladha, N., Agarwal, M., Majee, K. K., Das, A., Kumar, S., Saha, H. N. (2017). Low cost ultrasonic smart glasses for the blind. 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 210–213. https://doi.org/10.1109/ IEMCON.2017.8117194
- Aymaz, S., & Cavdar, T. (2016). Ultrasonic Assistive Headset for visually impaired people. 2016 39th International Conference on Telecommunications and Signal Processing (TSP), 388–391. https:// doi.org/10.1109/TSP.2016.7760903
- Blasch, E., Ling, H., Shen, D., Chen, G., Hammoud, R., Basharat, A., Collins. R., Aved, A &, Nagy, J. (2015).
 Video-to-text information fusion evaluation for level 5 user refinement. (2015). IEEE 18th International Conference on Information Fusion (Fusion): Retrieved from https://ieeexplore.ieee. org/document/7266799

- Demirovic, D., Skejic, E., & Serifovic-Trbalic, A. (2018). Performance of Some Image Processing Algorithms in Tensorflow. 25th International Conference on Systems, Signals and Image Processing (IWSSIP), 1–4. https://doi.org/10.1109/ IWSSIP.2018.8439714
- Gayathri, G., Vishnupriya, M., Nandhini, R., & Banupriya, M.M. (2015). Smart Walking Stick For Visually-Impaired. International Journal of Engineering and Computer Science IJECS. In International Journal of Engineering and Computer Science (Vol. 3). Retrieved from https://www.ijecs.in/index. php/ijecs/article/view/151
- GitHub navit-gps/navit: The open-source (GPL v2) turn-by-turn navigation software for many OS. (2018). Retrieved May 22, 2019, from https:// github.com/navit-gps/navit
- Hashino, S., & Ghurchian, R. (2010). A blind guidance system for street crossings based on ultrasonic sensors. The 2010 IEEE International Conference on Information and Automation, 476–481. https:// doi.org/10.1109/ICINFA.2010.5512383
- International Society of Information Fusion, IEEE Aerospace and Electronic Systems Society, & Institute of Electrical and Electronics Engineers. (n.d.). Information Fusion (Fusion), 2015 18th International Conference on : date, 6-9 July 2015. Retrieved from https://ieeexplore.ieee.org/ document/7266799
- Jose, A., George, G., Nair, M. R., Shilpa, M. J., & Mathai, M. B. (2016). Voice-enabled smart walking stick for visually impaired. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 5, 80-85.
- Kumar, K., Champaty, B., Uvanesh, K., Chachan, R., Pal, K., & Anis, A. (2014). Development of an ultrasonic cane as a navigation aid for the blind people. International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 475–479. https://doi. org/10.1109/ICCICCT. 2014.6993009
- Lakdawala, B., Khan, F., Khan, A., Tomar, Y., Gupta, R., & Shaikh, A. (2018). Voice to Text transcription using CMU Sphinx A mobile application for healthcare organization. *Second International Conference*

on Inventive Communication and Computational Technologies (ICICCT), 749–753. https://doi.org/10.1109/ICICCT.2018.8473305

- Nada, A., Fakhr, M. A, and Seddik, A. F. (2015). Assistive Infrared Sensor-Based Smart Stick for Blind People. SAI Conference- IEEE Technically Sponsored Science and Information Conference 2015, London, United Kingdom, 1149–1154. https://doi.org/10.1109/SAI.2015.7237289
- NAVIT running on a Raspberry Pi YouTube. (2018). Retrieved May 22, 2019, from https://www. youtube.com/watch?v=ldjWS0YIaMg
- Over 190,000 persons are blind in Ghana. (2016, October 13). Retrieved from https://www.ghanaweb. com/GhanaHomePage/NewsArchive/Over-190-000-persons-are-blind-in-Ghana-477283
- Prakoso, H., Ferdiana, R., & Hartanto, R. (2016). Indonesian Automatic Speech Recognition system using CMUSphinx toolkit and limited dataset. International Symposium on Electronics and Smart Devices (ISESD), 283–286. https:// doi.org/10.1109/ISESD.2016.7886734
- Sen, A., Sen, K., & Das, J. (2018). Ultrasonic Blind Stick for Completely Blind People to Avoid Any Kind of Obstacles. 2018 IEEE SENSORS, 1–4. https:// doi.org/10.1109/ICSENS.2018.8589680
- Shivakumar, K. M., Jain, V. V., & Priya, P. K. (2017). A study on impact of language model in improving the accuracy of speech to text conversion system. *International Conference on Communication and Signal Processing (ICCSP)*, 1148–1151. https:// doi.org/10.1109/ICCSP.2017.8286557
- TensorFlow How to Retrain Inception's Final Layer for New Categories Tutorial-YouTube. (2018). Retrieved May 22, 2019, from https://www. youtube.com/watch?v=m2D02eZTB4s
- Tudor, D., Dobrescu, L., & Dobrescu, D. (2015). Ultrasonic electronic system for blind people navigation. E-Health and Bioengineering Conference (EHB), 1–4. https://doi.org/10.1109/ EHB.2015.7391545