Computer Vision Based Indoor Navigation for Shopping Complexes

G.S.T Perera
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
Samadhi0830@gmail.com

K.W.R. Madhubhashini
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
raveesha96127@gmail.com

Dilani Lunugalage
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
dilani.l@sliit.lk

Piyathilaka D.V.S
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
sathipiyathilaka@gmail.com

Lakshani W.H. U
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
ureshalak97@gmail.com

Dharshana Kasthurirathna
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
dharshana.k@sliit.lk

ABSTRACT

Smartphone-based indoor navigation systems are frantically required in indoor situations. This limitation of clients is significant. Global Positioning System (GPS) isn't plausible for indoor areas as it gives exceptionally helpless outcomes for indoor restriction. In this research paper, we present a Computer Vision-Based Indoor Navigation System for shopping complexes. Computer vision is used in this system to find the exact location/current location of the user. It contains a mobile android application for positioning, navigating, and displaying the current location for showing on 2D Map. The system will detect the user's position, generate a GIS map, display the shortest path using A* search algorithm, and provides step-by-step direction to the destination using audio instruction for localization with Augmented Reality (AR) map and navigation using mobile phone sensor technologies like accelerometer, gyroscope, and magnetometer. The audio instructions include active guidance for upcoming turns in the traveling path, distance of each section between turns. This system uses a suggestion-based Chabot that uses a trained model to improve the user's experience. Thus, this research expects to build a cost-effective, efficient, and timely response system that will help the users for a smart shopping experience.

CCS Concepts

• **Computing methodologies → Artificial intelligence → Computer vision → Computer vision tasks → Activity recognition and understanding**

Keywords

Computer Vision, Convolution Neural Network, A* Algorithm, Augmented Reality (AR)

ICVISP 2020, December 9 - 11, 2020, Bangkok, Thailand

ACM ISBN 978-1-4503-8953-2/20/12…$15.00
https://doi.org/10.1145/3448823.3448828

ACM Reference format:

G.S.T Perera, K.W.R. Madhubhashini, Dilani Lunugalage, Piyathilaka D.V.S, Lakshani W.H.U and Dharshana Kasthurirathna. 2020. Computer Vision Based Indoor Navigation for Shopping Complexes. In Proceedings of *2020 4th International Conference on Vision, Image and Signal Processing(ICVISP 2020), December 9-11, 2020, Bangkok, Thailand*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3448823.3448828

1. INTRODUCTION

Shopping malls are places with lots of shops and people for shopping. With the advancement of technology, people try to complete the activities in less time. If the navigation system provides less accurate localization details, it affects the users. For the outdoor localization systems, GPS is used but it's not an appropriate way for the large indoor spaces due to the weakened GPS satellite signals, unable to detect floor numbers through the indoor environment because of the height of the building and the sensitivity of GPS to environmental conditions. The concept of indoor navigation systems has been introduced to overcome the issues of GPS technology for indoor spaces.

Computer Vision technology is used to detect the exact location of the user. In this proposed system convolutional neural network (CNN) is used. Nowadays convolutional neural network is widely used for object identification. The CNN model is trained by a training dataset. Since this system is based on machine learning (Speech Synthesis) and it provides more accurate and precise guidance to the user.

With the rapid development of technology, sensors are used for indoor localization systems. Wi-Fi fingerprinting sensors, these existing sensor-based systems are not popular among people because of security issues [13]. Lack of security is the main disadvantage and there will be hacks and security patches. Revolution of the smartphones has changed communication and everyone uses mobile phones as their main method of communication nowadays and mobile phone sensors like gyroscope, accelerometer, and magnetometers can be used in indoor navigation systems.

To provide a novel experience for shopping mall users, one of the features that we are going to concern about is sending audio instruction while shopping. A guided AR-based map consisting of audio navigation from a chatbot makes it easier to find where the user headed if they need suddenly to change the destination for a quick detour the chatbot created with the other systems helps the user to reach that desired destination from the current location by typing a simple question.

This is a supreme environment to explore novel applications. According to findings, there is no such web-based portal or mobile application to enhance customer satisfaction with 2D maps, computer vision technology, mobile phone sensor technology, and AR technology with audio instructions. The main aim of the system is to develop an Indoor Navigation system for shopping malls using computer Vision Technology with mobile phone sensor technology.

2. RELATED WORK

During the past years, there have been a lot of research projects proposed to develop an accurate solution for indoor navigation which has been proposed to users who are shopping complexes. When studying available resources, Aerospace Convergence Technology Team, Korea Aerospace Research Institute [8] proposed a design for large scale map generation which uses Simultaneous Localization and Mapping (SLAM) technologies that generates a 2D image of the map. Using Sensors such as LiDAR (Light Detection and Ranging), SONAR (Sound Navigation and Ranging), and RGB-D camera a 2D image can be generated with the help of Hector Simultaneous Localization and Mapping technology.

All in all, the Computer vision-based frameworks use a smartphone or a gadget installed with a camera, for example, a Google Glass to catch the scenes while the client is strolling through the indoor regions. Restriction utilizing PC vision innovation requires acknowledgment of the visual scene by coordinating the caught picture with existing pictures or by utilizing prepared models. Such models vary contingent upon the focused-on destinations, for example, support vector machine (SVM) model [14], neural organization model including profound learning model [15].

The main PC vision-based framework in particular 'CamNav' uses an indoor scene acknowledgment procedure to assess the situation of the client. CamNav utilized a prepared model to perceive the area from the question or caught pictures. The caught pictures are prepared in the worker utilizing a prepared profound learning model [9]. 'QR Nav' is a QR code based framework which is used for visual markers called QR codes glued in the indoor zones to decide the area of the client. The QR codes contained in the caught pictures are decoded by a condition of craftsmanship QR code decoder executed in the worker [10]. The BLE based framework comprises of an Android application and a foundation comprised of BLE reference points.

F. Serhan Danis¸ and Ali Taylan Cemgil from Department of Computer Engineering, Bogazici University, and Istanbul, Turkey discovered one of the factors is the gotten signal quality marker, which may give a gauge of the flight separation of the sign from the guide to the collector, after widely investigating the presentation of BLE innovation in indoor conditions. [2].

Beakcheol Jang, Hyunjung Kim, and Jong Wook Kim who is from the Department of Computer Science, Sangmyung University, Seoul, Korea proposed a framework called 'BatTracker' for indoor route framework [7]. 'BatTracker' utilizes sensors to gauge the client's position dependent on inertial information and rectifies blunders by estimating separation utilizing acoustic signs reflected off close by objects.

Chiaki Takahashi, Kazuhiro Kondo (2015), presented a practicality study and a framework to swap strategies for screen direction for the indoor route with the assistance of sound by utilizing iBeacons. The creators investigated how the area exactness can be accomplished by utilizing a predetermined number of iBeacons put inside indoor space and breaking down the radio gathering power from the signals [11]. The assessed position is the situation in the information base with the littlest separation with the deliberate power. The creators utilized four iBeacons for the analysis and had the option to accomplish 0.8m of position assessment mistake.

When comparing the existing indoor navigation systems with the newly implemented system, there are some limitations (table 1). This developed system is easy to find users' current locations using computer vision as well as AR technology and easy to navigate in shopping mall using sensor technologies. Although, Visual impairments persons' can use this system because it provides audio instructions for navigation and have a facility to questioning and answering (chatbot) using audio.

Table 1. Comparison of existing systems with the implemented system.

References	Technology	Techniques	Remarks
QRNav	Computer vision	QR Code Recognition	Requires huge measure of QR codes for safe route
BLE Base system	BLE	Fingerprinting & trilateration	Generally low precision brought about route blunders
Implemented system	Computer vision	CNN based image recognition Mobile phone sensors AR technology	Easy to find location. No need of external devices High in accuracy because of the values which are taken from mobile sensors

3. METHODOLOGY

The system that is proposed in this research consists of two applications which are a web application and a mobile application. The web application is mostly handled by the system admin or owner of a particular building to upload the floor plans of the building and configure the building as required.

The mobile application is built using android. This is targeted only for the users since it allows the users to navigate through the indoor environment using the shortest, most preferable, and popular path. Also helps in tracking the location of the user in an indoor environment.

Figure 1 and 2 shows that the component diagram and the high-level architectural diagram of the system.

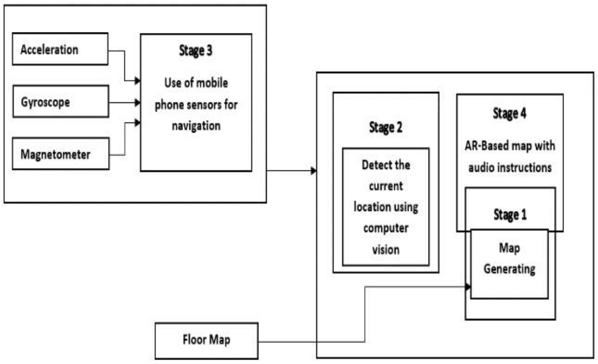

Figure 1. Component diagram.

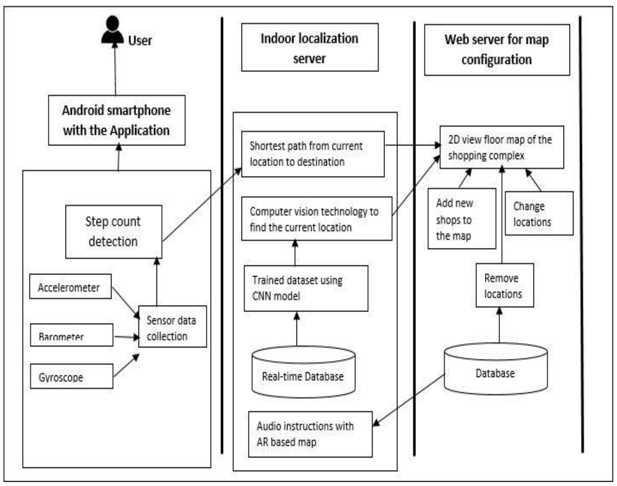

Figure 2. High-level architectural diagram.

The main functionalities of the Vision-based indoor navigation systems are GIS map generation, computer vision technology to find the users' current location, mobile phone sensor technology for navigation and AR based map with audio instructions.

3.1 GIS Map Generation

Geographic Information System mapping component is a GIS map generation using a 2D view map which helps users to navigate the path to the destination by looking at a map through an indoor environment. From GIS map generation, a 2D map will be generated while the other functions will work on top of the map. After the user finds the user's exact location in the indoor environment and points it to the user the indoor 2D view map is displayed on the user's mobile device. The basic idea of the GIS mapping in the system is shown in figure 3. The objective is a GIS mapping component to create a 2D GIS map on any given floor plan which has been created with the requested and predefined standards. This standard had been assigned to the floor plan to get more accurate readings on the floor plan [6].

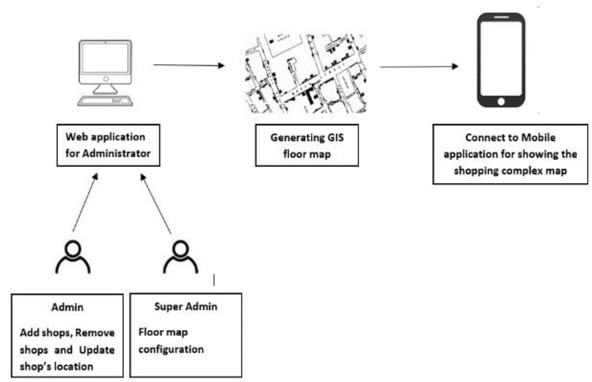

Figure 3. The basic principle of web GIS.

3.2 Computer Vision Technology to Find The Users' Current Location

Computer vision technology is used to detect the exact location of the user or update the user's current location. In this functionality, Restriction utilizing computer vision innovation requires acknowledgment of the visual scene by coordinating the caught picture with existing pictures or by utilizing prepared models (CNN Model). Computer Vision Technology uses signboards' recognition to estimate the location on the map whenever a known sign is recognized. For this functionality, Use the Convolution Neural Network (CNN) model to train the image data set which is shopping malls each shop signboard [3].

When discussing image classification functionality, CNN (Convolution Neural Network) model used to identify shop signboard images because of many reasons. Such as, CNN compares the image piece by piece. So, it gets better at seeing similarities than whole image matching schemes. It can be thought of programmed include extractors from the picture and utilizes nearby pixel data to viably down example the picture first by convolution and afterward utilizes an expectation layer toward the end. It performs well feature extraction and it gives better accuracy.

The server application uses a prepared deep learning model to perceive the area of the picture. It returns the location information to the mobile application in a JSON format.

Main steps of (When the user enters the shopping mall)

- Capture the nearest shop's signboard (user)
- Upload signboard image to the system (user)
- The system identifies the exact shop using a trained dataset
- The system shows the shops located on the floor 2D map (user's current location)

In this functionality, the outcome is the Users' current location on the 2D floor map. Figure 4 shows the steps of users' current location detection using computer vision technology.

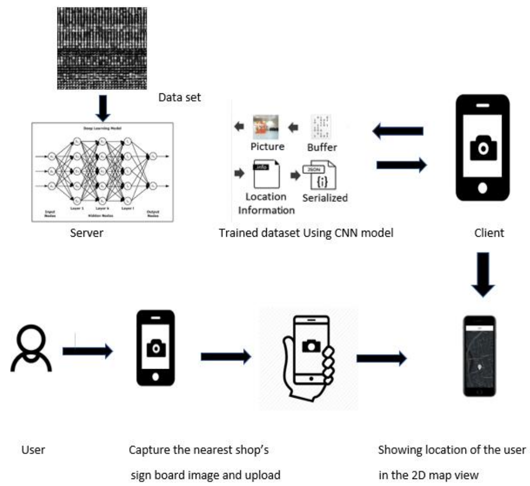

Figure 4. Computer vision technology to detect users' exact location component.

3.3 Mobile Phone Sensor Technology for Navigation

Navigation to the destination using sensor technology is one of the main components of the indoor navigation system. Through sensor technology, users will be able to see the shortest path or the most preferable path for the destination and the users can be identified to the points where they have to turn until their destination. Values that are taken from sensors (Accelerometer, gyroscope, magnetometer, and barometer) are used to create an algorithm for finding the shortest path or most suitable path to the destination from the current location of the user. Figure 5 shows the overview of the sensor technology for navigating.

Mainly three functions use sensor data values for the calculations. Step count detection, user's walking distance calculation and user-directed side detection are those three functions [4].

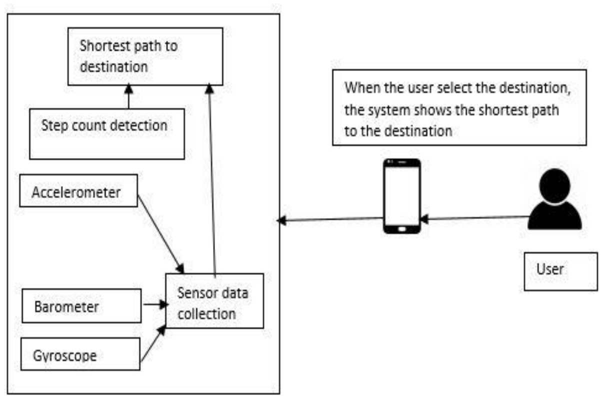

Figure 5. Overview of sensor technology for navigating.

3.3.1 Step Count Detection

The regular technique to identify a stage is to distinguish the pinnacles of vertical increasing speed, which compare to the progression events because the vertical quickening is created by the vertical effect when the foot hits the ground [12]. At the point when the framework distinguishes the progression of a client, it figures two qualities for the separation and the heading bearing [1]. In time t, the framework figures separation D (t) utilizing the estimations of the speeding up sensor α (t) as indicated by the beneath condition [5].

$$D(t) = \int_{t-1}^{t} \left(\int_{t-1}^{t} \alpha(t) D(t) \right) D(t) \qquad (1)$$

The main focus of the research is to get the shortest path and the user's preferable path in a short period and use the simplest way to get the most accurate solution. A* algorithm is a searching algorithm that searches the shortest path between the initial node and the final node. As we are using a 2-D indoor map, the A* algorithm is the most appropriate algorithm to find the shortest and the most preferable path. Therefore, among these algorithms, the A* search algorithm is one of the best and popular techniques used in 2-D maps.

3.4 AR Based Map With Audio Instructions and Chatbot

For navigation this system uses AR technology and the users will navigate throughout the indoor environment using 3D shapes rendered in front of the camera view, by extract real-time orientation from both accelerator and magnetometer the system will detect the user's current position and will identify navigation instruction to be given at that location. For the identified location navigation shapes will be rendered and following a series of rendered navigation shapes users will be guided to the destination.

AR Core utilizes three key abilities to incorporate virtual substance with this present reality as observed through the telephone's camera. [9]

By giving the guidance and voice orders and bolts on the guide. The System ought to have the option to answer to the clients in all circumstances and the figure 6 presents the process of the audio instructions for navigating with an AR based map and chatbot. The particularly framework needs to utilize voice answers to the client to discover the way to the objective, the proposed framework utilizes TensorFlow to make a neural organization and train it with the expectation document to create reaction model. This reaction model can be utilized to anticipate the reaction from the question of the client.

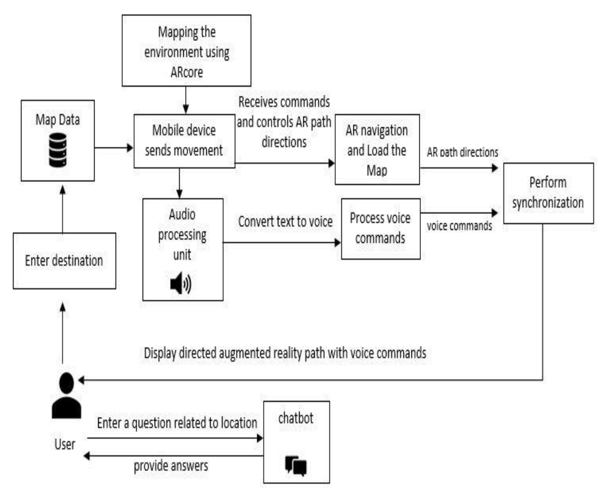

Figure 6. The process of the audio instructions for navigating with an AR-based map and chatbot.

4. RESULTS AND DISCUSSIONS

A mobile application is targeted mainly for the users who come to the shopping complex and they can track their current location in the building by using a 2D view map and can navigate to any location inside the indoor environment also they can get an experience of AR-based map with audio instructions of the directions in an indoor environment.

We assessed the exhibition of a vision-based indoor navigation system in a certifiable domain. The assessment tests were completed on the first floor and the second floor of the Majestic city (MC) Building in Colombo. We chose 10 individuals including 6 females and 4 guys to assess the route frameworks progressively. The period of members is gone from 22 to 39 (Mean = 29.30, standard deviation (SD) = 4.90).

The members were approached to stroll from the passage entryway of the Majestic city (MC) working to two explicit focal points in the MC building. Every member needs to stroll from direct A toward B, (separation = 30 m) and A to C (separation = 47 m) utilizing the vision-based route frameworks independently. We recorded all the strolling tests utilizing a camcorder for additional evaluations and after that, they should give feedback (ratings 1-5) for the application.

For vision-based frameworks, a Samsung Galaxy S7 smartphone was utilized for running the android application. The versatile information of the cell phone was used as a vehicle for sending the pictures from the Android application to the server. On the server-side, an Asus PC with 24 GB smash was used for preparing the caught Shops signboard pictures. In these frameworks, we have noticed that the outcomes are influenced by the strolling space of the client. This is because when the client is strolling at a higher speed, the possibility of getting a hazy signboard image is higher.

Figure 7 shows that the accuracy rate of the CNN model's image classification.

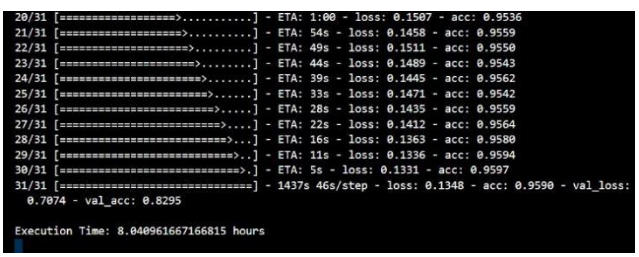

Figure 7. Accuracy rate for image classification.

It tends to be seen that after 31 ages the approval precision is 0.9590, it shows the capacity of the model, to sum up to new information. Because of testing the model, got an excellent precision: 95% of right characterization tests after 31 ages. The main disadvantage was that needed to hold up around 8 hours until 31 ages arrive at the end.

The table 2 shows that the error rate of 50 steps counts details for navigation.

Error tare is calculated by using the below equation.

$$(|\,approx - exact\,|/exact) \text{ *100\%} \qquad (2)$$

Table 2. Error rate for 50 step counts.

Test number	Step count	Error rate
1	50	0
2	52	4
3	50	0
4	51	2
5	50	6
6	51	0
7	50	0
8	50	0
9	52	2
10	50	0
Average	60.6	1.4

To inspect the viability of these frameworks, we arranged a poll for the individuals who partook in the investigations. In the wake of finishing the ongoing tests, we requested that the members round out the survey. In the survey contains, the members are approached to rate the fundamental functionalities (Location assessment utilizing billboard pictures, Navigation way proposal, and Speech directions) of the route frameworks utilizing a Likert size of focuses 1 (terrible) to 5 (superb).

After gathering clients' criticism for this indoor navigation framework, it appears to be acceptable appraisals for it. All users gave great evaluations in the wake of utilizing a vision-based indoor route framework for their route in that shopping complex. We suggested a vision-based indoor route application as it is generally appropriate for indoor zones. Clients can have a better shopping experience utilizing this framework.

5. CONCLUSION AND FUTURE WORKS

In this article, we created a vision-based indoor navigation system for individuals who visit shopping centers for their requirements. The proposed frameworks use CNN based scene acknowledgment, situating to explore individuals in the shopping center. The frameworks were tried on individuals in a continuous situation. In ease of use analyzes, this framework got the most elevated framework convenience score. Implanting an impediment discovery utilizing computer vision innovation, sensor innovation, AR innovation, and sound directions alongside route frameworks can guarantee the security of the client during the presence of a portable snag.

Later on, we will consider incorporating QR code acknowledgment in a vision-based framework to adapt to the ID of the area which is comparable in appearance. The future variant of the framework will address the majority of the prerequisites recommended by the clients who took an interest in the constant assessment in shopping centers. The estimation of the step count can be derived by the strength of the acceleration movements of the accelerometer. In future work, we are planning to expand the application for exhibitions as a commercial application with an automation process. We planned to develop remote shopping, personalized recommendations, and get to know further information about a particular product for people to get the same smart shopping experience as other customers as well, and we planned to develop an automatic bill calculating part for the shopping as well.

ACKNOWLEDGEMENT

A vision-based indoor navigation system for shopping complexes Mobile Application Research was carried out as the 4th year research project in Sri Lanka Institute of Information Technology, Malabe, Sri Lanka. We are extremely grateful to the Department of Administration of Majestic City Colombo for providing the related resources which helped to the research.

REFERENCES

[1] Chen, T. and Xu, L. (2015). *International Conference of Electrical, Automation and Mechanical Engineering*. An Improved Dead Reckoning Algorithm for Indoor Positioning Based on Inertial Sensors. China: Z. Zhou, p.371.

[2] Danis, S. and Cemgil, A. (2017). *Model-Based Localization and Tracking Using Bluetooth Low-Energy Beacons*. Turkey, pp.1-23.

[3] Jayakanth Kunhoth, AbdelGhani Karkar ,Somaya Al-Maadeed ,Asma Al-Attiyah, Comparative analysis of computer-vision and BLE technology based indoor navigation systems for people with visual impairments, 2019.

[4] Li, F., Zhao, C., Ding, G., Gong, J., Liu, C. and Zhao, F. (2012). *A Reliable and Accurate Indoor Localization Method UsingPhone Inertial Sensors*. China.

[5] Lu Jingge, yang Yongguo, ",WebGIS server based on open source software to build[J]. GEOMATICS & SPATIAL INFORMATION TECHNOLOGY", 20088, 31(5):pp.145-147.

[6] Luo, C., Cheng, L., Hong, H., Sankaran, K., Chan, M., Li.J. And Ming.Z.(2018) From mapping to indoor semantic queries: Enabling zero-effort indoor environmental sensing. [Online] Available at: https://www.sciencedirect.com/science/article/pii/S1084804516302910 [Accessed 23 Mar.2018].

[7] Marotto, V., Serra, A., Carboni, D., Sole, M., Dessi, T. and Manchinu, A. (2020). *The Fourth International Conference on Sensor Device Technologies and Applications*. Orientation Analysis through a Gyroscope Sensor for Indoor Navigation Systems. Italy, p.85.

[8] Moon, S., Ecom, W. and Gong, H. (2018). Development of Large-scale 3D Map Generation System for Indoor Autonomous Navigation Flight – Work in Progress. [Online] Available at: https://www.sciencedirect.com/science/article/pii/S1877705814038119 [Accessed 24 Mar.2018].

[9] S.A Kumarajeewa, P.I. Bodaragama, Nethmi Lokubalasuriya and P.D. Sandaruwan (2010) Smart Voice Recognition Based Navigation System - Sri Lanka Institute of Information Technology.

[10] Shenlong Wang, Sanja Fidler, Raquel Urtasun: LostShopping! MonocularLocalizationinLargeIndoorSpaces.

[11] Takahashi, C., & Kondo, K. (2015). Indoor positioning method for augmented audio reality navigation systems using iBeacons. 2015 IEEE 4th Global Conference on Consumer Electronics (GCCE). doi:10.1109/gcce.2015.7398636 https://ieeexplore.ieee.org/document/7398636

[12] W. C. S. S. Simões and V. F. de Lucena, "Blind user wearable audio assistance for indoor navigation based on visual markers and ultrasonic obstacle detection," 2016 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, 2016, pp. 60-63. https://ieeexplore.ieee.org/document/7430522

[13] Xia, S., Liu, Y., Yuan, G., Zhu, M. and Wang, Z. (2017). Indoor Fingerprint Positioning Based on Wi-Fi: An Overview. 6(135), p.1.

[14] Yin H, Jiao X, Chai Y, Fang B. Scene classification based on single-layer SAE and SVM. Expert Syst Appl. 2015;42(7):3368–80.

[15] Zhou B, Lapedriza A, Khosla A, Oliva A, Torralba A. Places: A 10 million image database for scene recognition. IEEE Trans Pattern Anal Mach Intell. 2017; 40(6):1452–64.