Vision Based Office Assistant Robot System for Indoor Office Environment

S. I. A. P. Diddeniya
Department of Physics
University of Sri Jayewardenepura
Nugegoda, Sri lanka.
isankadiddeniya@sjp.ac.lk

A. M. S. B. Adikari
Department of Physics
University of Sri Jayewardenepura
Nugegoda, Sri lanka.

H. N. Gunasinghe
Department of Computing and Information Systems
Sabaragamuwa University of Sri Lanka
Belihuloya, Sri Lanka.

P. R. S. De Silva
Department of Computer Sciences
University of Sri Jayewardenepura
Nugegoda, Sri lanka.

N. C. Ganegoda
Department of Mathematics
University of Sri Jayewardenepura
Nugegoda, Sri lanka.

W. K. I. L. Wanniarachchi*
Department of Physics
University of Sri Jayewardenepura
Nugegoda, Sri lanka.
iwanni@sjp.ac.lk

***Abstract*—This paper presents an office assistant robot that can be used in an unstructured indoor office environment. Among many technologies available, we used free and open source software and inexpensive sensors and materials to build the low cost but accurate robot. Robotic Operating System (ROS) indigo was used as the ground operating system on Ubuntu 14.04. The mobile robot, iRobot Create 2 was used as the basic robot and a structure was built to carry a mini-laptop and PrimeSense 3D vision sensor. A workstation computer was used as the central location PC which was kept still and map of office environment is built on it. User is able to give commands to the system via voice and virtual keys by the developed Android application (App). These three units, mobile robot, central workstation and Android devices, communicates through a Wi-Fi connection. The proposed robot could deliver documents or parcels in between office members according to the user commands. We allowed the robot to navigate autonomously and randomly between users and monitored its accuracy by looking at the completion of the route to a target user. Results show that the office assistant delivery robot has above 92% of accuracy in delivery process for a valid user input.**

***Index Terms*—Service Robots, RGB-D Perception, SLAM, Localization, Mapping, Visual-Based Navigation, Voice Recognition**

I. Introduction

Robots have become an emerging part in the world now because they help people to do tasks easily. There are several types of robots. Industrial robots, Domestic or household robots, Medical robots, Service robots, Military robots, Entertainment robots, Space robots, Hobby and competition robots are few examples for robot types. A service robot could assist human beings by performing a job that is delivery things, cleaning, and including any other household chores. Some service robots need to move place to place in order to complete their tasks. TurtleBot [1], the delivery robot SaviOne [2] and ASIMO [3] robots are few famous examples. Therefore, robot navigation is a very important part of delivery service robots. Humans could navigate in between known two places autonomously. In human autonomous navigation, the environment is captured by human eyes, and human brain identifies the path according to the captured environment data to move to the correct location. A vision based robot autonomous navigation system imitates the human autonomous navigation. In an autonomous navigation system, the robot has a vision sensor work as eyes, movable robot base with controllable actuators that represent legs or wheels and processing unit that acts as the human brain. Also, delivery robot needs commands to move place to place to deliver goods between people.

Document delivery is a routine task in an office environment but it is very dull task for people. Hence, working time of office members is wasted continually. Also productivity of the office is reducing. Therefore we decided to introduce an office assistant robot for an indoor office to help people to sit back and do their daily work easily. Where by, office members can work in the office without wasting their working time and the productivity of the office is increasing.

In this research work, we implemented a low cost autonomously navigable office assistant robot for an indoor office environment. It can deliver a document or a small parcel ($<= 1kg$) in between two office members autonomously. In this proposed robot system, we used iRobot Create 2 programmable robot [4, 5] base which is a low-cost robot base, PrimeSense 3D sensor [6, 7] as the vision sensor and a mini-laptop computer as the processing unit of the robot. A personal computer was used as the center location PC. Further, Android application was developed for user input which enables to control the system with any android device. Every office member can install this Android application and give commands to robot to deliver their documents. Users can give their commands using voice and virtual key pressing on Android application. These units communicate via a private Wi-Fi network.

Fig. 1. Mobile Robot System: iRobot Create 2 base, PrimeSence 3D sensor and mini-laptop.

II. Related Work

On their research work, G. Chauhan and P. Chaudhari [8], a voice controlled mobile robot system was introduced. in there system, speech recognition was carried out in a separate desktop computer and mobile robot communicates with the PC via Bluetooth technology. Further, an Android app for keyboard controlling of the robot, was also implemented. In our proposed system, we control our robot using both methods; voice control and key pressing using an Android application through Wi-Fi. This communication is faster and covers wider range than Bluetooth. Therefore, our system could be used in large indoor areas and controlled effectively and efficiently.

In three separate researches, Y. A. Memon et al [9], S. Pinjarkar et al [10] and H. Rashid et al [11] introduced Android based robot controlling systems. They used an Android smart phone as a speech recognizer platform where online based voice recognition service utilized to control the robot. Here, the major drawback is the developed systems in their work relies on the internet connection. In contrast our proposed system used offline based Google voice recognition system (Default android voice recognition features) to control our robot. Further, we considered its English (United Kingdom) language database for Google voice recognition as proven by our own previous experiment [12].

Above mentioned robot systems mainly navigate in the environment according to the voice or keyboard commands given by the user. When it comes to autonomous navigation, there are several methods of robotic navigation. Ultra sound based [13], Laser based [14], GPS based [15] and Vision based navigation [16] are few examples. Those methods make use of laser, sonar, GPS, IMU, compass, RGB camera and 3D vision cameras as sensors. In the vision based robotic navigation, the robot needs a map to navigate and to know location of itself in the environment [17]. Therefore, depth value is very important to navigation and obstacle avoiding of robot. In vision based navigation system, RGB camera or 3D vision sensor is used as the main sensor. 3D vision sensor can capture RGB values with depth values (RGB-D) of object in the environment. Microsoft Kinect sensor, Apple PrimeSense, Intel RealSense and Asus Xtion Pro are most recent 3D vision sensors in consumer market. 3D vision sensors are being used in various fields including Home Automation Systems [18, 19] and Virtual Dressing Rooms [20].

J. Juang and C. Yang [21] proposed a document delivery robot based on image processing and fuzzy control system. In their developed system, they used RGB camera to capture the office environment and compiled with image processing techniques. This system could be used in light illumination conditions, but could not use in dark illumination conditions. We use 3D vision Sensor in our system. As we have addressed in [22] our previous test results with the proposed system, show that it can be used in both dark and light illumination conditions.

M. Jalobeanu et al [23] proposed a reliable Kinect-based navigation in large indoor environments. Here, they used Kinect RGB-D sensor, Pioneer 3DX robot platform, and LIDAR to build the robot system. LIDAR is more accurate sensor to build environment map for robot navigation, but very expensive. Therefore, our proposed system could be introduced as an inexpensive delivery robot system as we used the PrimeSense RGB-D sensor instead of LIDAR.

D. Sales et al [24] proposed a mobile robot navigation in indoor environments using Kinect sensor. They used Pioneer P3-AT robot as base of the robot and Kinect sensor as the 3D vision sensor. In order to satisfy the power requirement for the Kinect sensor, need extra power source (12 V, 2.67A). In addition, to interface Kinect to a computer need its own USB converter. Altogether, the mobile robot has to carry extra weight (cables, USB convertor, and battery pack) when a system implemented with a Kinect sensor. Further, Pioneer P3-AT robot base is more expensive than iRobot Create 2. Therefore, our proposed system is cost effective. Also their system is based on a finite state machine (FSM) learned by an artificial neural network (ANN) in an indoor navigation task. In that method, the robot cannot do localization during the navigation. In our proposed system, we used Simultaneous Localization and Mapping [25] (SLAM) algorithm for building the map and localization during the navigation.

K. Yamazaki et al proposed home assistant robot for an aging society and it can process few household tasks. It is a life size humanoid robot system. In their approach, they use a 3D geometrical simulator that defines the virtual world in which robots and objects are arranged mimicking the real world. This is very advanced for a service robot system but they mentioned that their system is not acceptable to stop the execution of the task when a failure is detected [26]. Also this system is very expensive. In our system, we planned inexpensive and easy

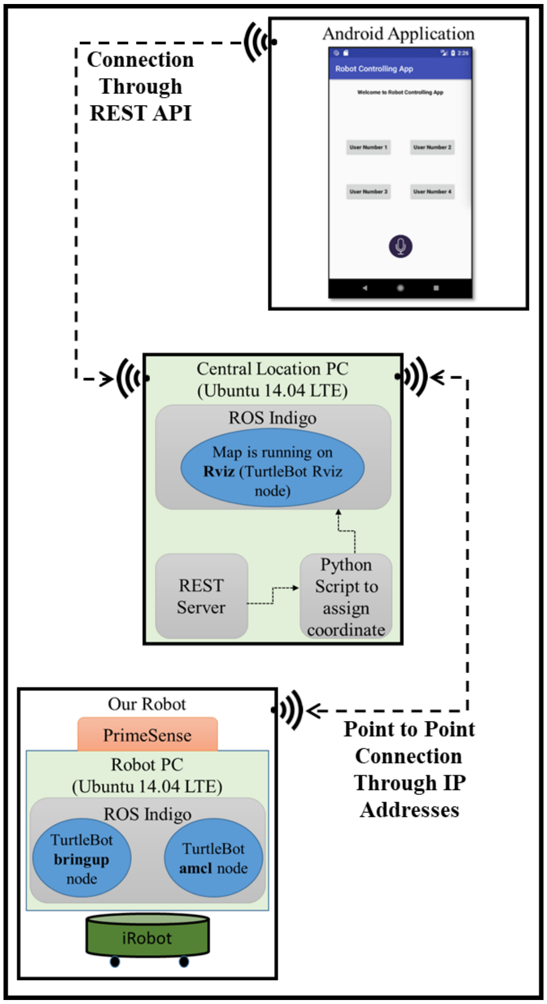

Fig. 2. System Overview.

handling office assistant robot system for indoor office.

III. Implementation

The developed office assistant robot system consists with main three units, such as mobile robot, centrol location PC and Android devices. The structure of the mobile robot was built using iRobot create 2 as the base. Primesense 3D vision sensor (USB powered) was used as the vision sensor. A mini-laptop (Intel Celeron Dual Core N3060 with DDR3 2GB RAM) used as the processing unit of the mobile robot which interface robot platform and the depth sensor while communicating with the central location PC (Figure 1) via Wi-Fi. A desktop computer (Intel Core i7 processors with DDR3 8GB RAM) was used for the central location PC. Android application was developed (Figure 2) for the user voice and key inputs which allows users to deliver documents to the desired locations. ROS (Robotic Operating System) indigo was installed on both mobile processor unit and the central location PC [27] on top of the Ubuntu 14.04 operating system. Turtlebot which is an existing package for ROS was selected as the robot kit. Among its numerous functionality we made the use of bring up, gmapping, keyboard teleop, Rviz view, and amcl (Adaptive Monte Carlo localization approach) nodes in this study.

A. Creating a map of environment (Mapping)

Robot uses a map to navigate autonomously in the environment. First, bring up, gmapping, keyboard teleop and Rviz view nodes were used to generate the map of the environment on the robot. The map is built on the XY Cartesian plain as shown in figure 3.A. To enable mapping process, bring up and gmapping nodes were executed on robot PC and keyboard teleop and Rviz view nodes were executed on Central Location PC.

In here SLAM algorithm [25] was used to implement the gmapping node. SLAM algorithm was introduced by J. J. Leonard and H. F. Durrant-Whyte [28] as the application of the extended Kalman filter (EKF). They developed an algorithm for model-based localization that relies on the concept of a geometric beacon a naturally occurring environment feature that can be reliably observed in successive sensor measurements and can be accurately described in terms of a concise geometric parameterization. The Kalman filter relies on two models: a plant model and a measurement model. The plant model describes how the vehicles position x(k) changes with time in response to a control input u(k) and a noise disturbance v(k).

$$x(k+1) = F(x(k), u(k)) + v(k), \qquad v(k) \sim N(0, Q(k))$$

Where $F(x(k), u(k))$ is the nonlinear state transition function. the used the notation $u(k) \sim N(0, Q(k))$ to indicate that this noise source is assumed to be zero mean Gaussian with variance Q(k). The measurement model expresses a sensor observation in terms of the vehicle position and the geometry of the beacon being observed and has the form,

$$z_{\mathrm{j}}(k) = h_{\mathrm{i}}(p_{\mathrm{i}}, x(k)) + w_{\mathrm{j}}(k), \qquad w_{\mathrm{j}}(k) \sim N(0, R_{\mathrm{j}}(k))$$

The observation function $h_i(p_i, x(k))$ expresses an observed measurement $z_j(k)$ as a function of the vehicle location x(k) and beacon location p_i. This observation is assumed corrupted by a zero-mean Gaussian noise disturbance $w_j(k)$ with variance $R_j(k)$. The form of the observation function $h_i(. , .)$ depends on the sensor employed and the type of beacon being observed. The goal of the cyclic computation is to produce an estimate of the location of the robot $\hat{x}(k+1 \mid k+1)$ (The term $\hat{x}(i \mid j)$ should be read as "the estimate of the vector x at time step i given all observations up to time step j") at time step k + 1 based on the estimate of the location $\hat{x}(k \mid k)$ at time step k, the control input u(k) and the new beacon observations $z_j(k + 1)$. The algorithm employs the following steps: prediction, observation, matching, and estimation.

B. Navigation on known map

Bring up, amcl, and Rviz view nodes and the generated map were used to autonomous navigation. To enable navigation process, bring up and amcl nodes were executed on robot PC and Rviz view node was executed on central location PC. In order to initialize the robot navigation, user has to

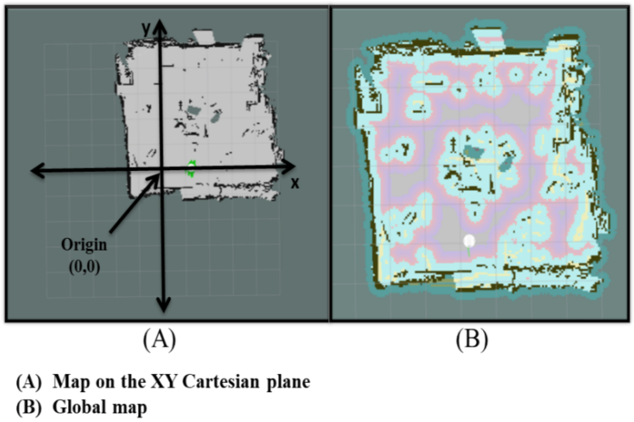

Fig. 3. Maps of the unstructured room.

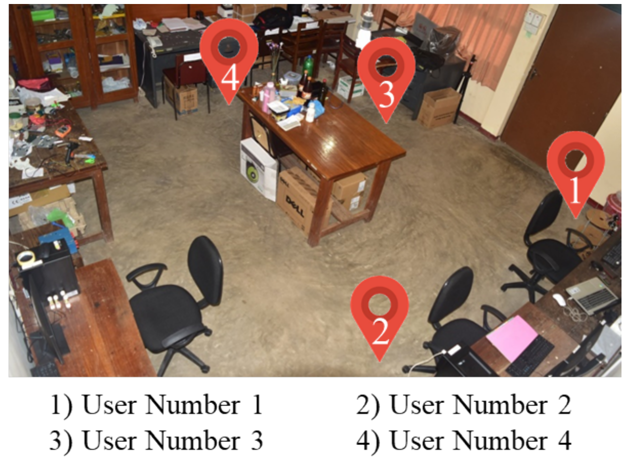

Fig. 4. Unstructured indoor environment of the system tested.

initiate a pose estimation which enables the robot to know its position on the map. The user has to provide a goal location using android application. Location is provided through a pre-assigned coordinates of user positions on map. Our robot calculates the minimum route to the goal location using the developed map in the first step. Then robot navigates to the given location. While its traverse, the robot updates the map and localize itself simultaneously according to SLAM algorithm. An illustration is shown in figure 3.B. As a result, the path gets updated.

This process based on Adaptive Monte Carlo Localization (AMCL) introduce by ROS according to the KLD-Sampling: Adaptive Particle Filters and Mobile Robot Localization research work of D. Fox [29]. It is a statistical approach for increasing the efficiency of particle filters by adapting the size of sample sets on-the-fly. The key idea of the KLD-sampling method is to bound the approximation error introduced by the sample based representation of the particle filter.

C. Android Application Overview

User is able to provide commands using either voice or virtual keys on Graphical User Interface (GUI) application. The GUI is designed to input the name of the expected recipient of the document. We have implemented a virtual button to enable the user to speak up and used Google Voice Input function [30] to receive the voice command. It is recognized using Google Voice Recognition.Also we have included name buttons in the GUI with previously assigned user names. Users can select desired recipient from the list of buttons. The string containing the name is sent to the central location PC. Using a Python script, the PC then searches for the coordinates of the recipient from a predefined list. Then those coordinates and the path to the goal which is obtained through the map are sent to the robot. Finally the robot moves accordingly.

D. Communication process of the system

ROS requires bidirectional network between the robot and central location PC attached to the private Wi-Fi network. The network for the robot and central location PC can be configured with their IP addresses. In order to achieve a communication, users' android devices are required to connect with the same network. When a command received at the android application, it passes the command to central location PC, acting as a Representational State Transfer [31] (REST) client. A python script is written in central location PC to process it being a RESTful web server.

IV. Experiment and Results

Our experiment was done in an unstructured room of dimension 5.7m × 5.6m × 3.3m. Accordingly the developed system was tested with Google voice typing for English (United Kingdom) language. We tested our system for four Sri Lankan native Sinhala language speaking users, two males and two females in order to increase the systems generalization. They are located to the sides of the unstructured room and named each user number 1 to 4 and assigned coordinates on map as shown in the figure 4. We tested the accuracy of the robot for user commands provided through the Android application using both voice and virtual keys.

First, the robot is placed on a random location inside the room. A user is given the smart phone with our application installed. Then, user was asked to call the robot to his/her location. Secondly, he/she has to provide recipient to deliver the document. We examined whether the robot act accordingly for each scenario. We monitored the success rate of completion of document delivery process for each user by conducting the experiment randomly for 20 times per user in each commanding methods.

The experimental result shows more than 92.5% accuracy on average. It could be seen that 90% of accuracy for achieving goal location of the office assistant robot and the system achievement has failed 8 times a total of 80 inputs of voice commands. Further, the robot has achieved the goal location 95% accuracy in 80 total rounds in virtual key pressing. Results are shown in the figure 5.

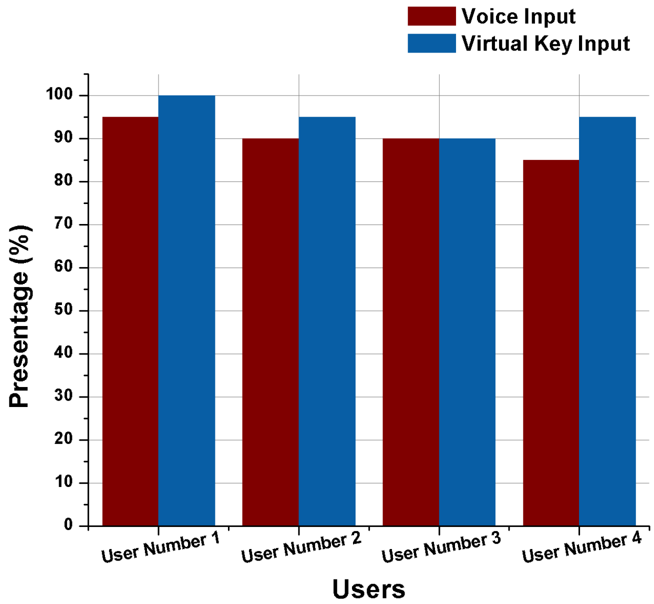

Fig. 5. System Performance in unstructured indoor enviroment.

V. Discussion

The proposed system has tested with native Sri Lankan users with Sinhala personal names, where the pronunciation is different than English names. As we have used Google Voice Typing with British English Voice Database, we considered technical names such as user number 1 instead of actual names of the users when testing. It is a drawback of this work that we are planning to address in future.

In the experiment, voice command was not recognized at 3 times, because Google Voice Typing was failed at those times to recognize correctly. Also robot was failed 9 times to achieve goal target, because robot localization and path identification were not occurred correctly. We are planning to access these problems immediately.

The developed robot system will be further improved in the near future to identify correct recipient using face recognition features. Finally, we hope to implement an office assistant robot for an indoor office environment.

VI. Conclusions

Localization and mapping are the two main features of a movable robot. In order to navigate safely towards a target locations, avoiding obstacles in an unstructured environment is a challenging task. We have built an autonomous office assistant robot system having two hardware units and wireless Android application. A separate processing unit and a movable robot structure were used in advance. These two components comprised of ROS indigo, PrimeSense and iRobot Create 2 that communicated in between. Using Android application, user can operate delivery robot using voice command and pressing virtual keys. Android application communicates with central location PC using REST APIs through Private Wi-Fi connection. Our system was trained in an unstructured indoor environment and tested its accuracy of the delivery process according to the user input voice commands and virtual key pressing commands using four users. The experiment results prove that the implemented system has high accuracy and it can be successfully used in an unstructured environment and the robot will help to make users delivery tasks easier.

Acknowledgment

The authors would like to thank the University of Sri Jayewardenepura (Grant No ASP/01/RE/SCI/2017/13) for th efunding support and to Mr. O. K. D. M. Priyantha, Workshop Technical Officer, Faculty of Applied Sciences, University of Sri Jayewardenepura for the given support.

References

[1] Hesham Ibrahim Mohamed Ahmed Omara and Khairul Salleh Mohamed Sahari. Indoor mapping using kinect and ros. In *Agents, Multi-Agent Systems and Robotics (ISAMSR), 2015 International Symposium on*, pages 110–116. IEEE, 2015.

[2] Neha Shukla. Savione the butler bot: Service robot for hospitality industry. *TechieTonics, http://www. techietonics. com/robo-tonics/savione-the-butler-bot-service-for-hospitality-industry. html*, pages 1–5, 2014.

[3] Yoshiaki Sakagami, Ryujin Watanabe, Chiaki Aoyama, Shinichi Matsunaga, Nobuo Higaki, and Kikuo Fujimura. The intelligent asimo: System overview and integration. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, volume 3, pages 2478–2483. IEEE, 2002.

[4] Wolfgang Hönig, Arash Tavakoli, and Nora Ayanian. Seamless robot simulation integration for education: A case study. In *Workshop on the role of simulation in robot programming at SIMPAR*, 2016.

[5] iRobot Corporation. irobot® create® 2 open interface (oi) : Specification based on the irobot® roomba® 600.

[6] Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren, and Dieter Fox. Rgb-d mapping: Using depth cameras for dense 3d modeling of indoor environments. In *In the 12th International Symposium on Experimental Robotics (ISER*. Citeseer, 2010.

[7] Barak Freedman, Alexander Shpunt, Meir Machline, and Yoel Arieli. Depth mapping using projected patterns, April 3 2012. US Patent 8,150,142.

[8] Gaurav Chauhan and Prasad Chaudhari. Robotic control using speech recognition and android. *International Journal of Engineering Research and General Science*, 3(1):1210–1216, 2015.

[9] Y. A. Memon, Imaaduddin Motan, M. A. Akbar, Sarmad Hameed, and M. U. Hasan. Speech recognition system for a voice controlled robot with real time obstacle detection and avoidance. *International Journal Of Electrical, Electronics And Data Communication*, 4(9):33–37, 2016.

[10] Sagar Pinjarkar, Siddhi Khadpe, Anuja Tavte, and Riddhi Karpe. Voice controlled robot through android applica-

tion. *International Research Journal of Engineering and Technology (IRJET)*, 4(04), 2017.

[11] Humayun Rashid, Iftekhar Uddin Ahmed, Sayed Bin Osman, Qader Newaz, Md Rasheduzzaman, and SM Taslim Reza. Design and implementation of a voice controlled robot with human interaction ability.

[12] SIAP Diddeniya, KASV Rathnayake, HN Gunasinghe, and WKIL Wanniarachchi. Integrating irobot-arduino-kinect systems: An educational voice and keyboard operated wireless mobile robot. *International Journal of Scientific Engineering and Technology*, 6(6):188–192, 2017.

[13] Alberto Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal on Robotics and Automation*, 3(3):249–265, 1987.

[14] Karlan D Boultinghouse, J Lee Schoeneman, and Bertice L Tise. Method for laser-based two-dimensional navigation system in a structured environment, January 3 1989. US Patent 4,796,198.

[15] Giulio Reina, Andres Vargas, Keiji Nagatani, and Kazuya Yoshida. Adaptive kalman filtering for gps-based mobile robot localization. In *Safety, Security and Rescue Robotics, 2007. SSRR 2007. IEEE International Workshop on*, pages 1–6. IEEE, 2007.

[16] I Ohya, Akio Kosaka, and Avi Kak. Vision-based navigation by a mobile robot with obstacle avoidance using single-camera vision and ultrasonic sensing. *IEEE Transactions on Robotics and Automation*, 14(6):969–978, 1998.

[17] Stephen Se, David Lowe, and Jim Little. Vision-based mobile robot localization and mapping using scale-invariant features. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, volume 2, pages 2051–2058. IEEE, 2001.

[18] KASV Rathnayake, SIAP Diddeniya, WKIL Wanniarachchi, WHKP Nanayakkara, and HN Gunasinghe. Voice operated home automation system based on kinect sensor. In *Information and Automation for Sustainability (ICIAfS), 2016 IEEE International Conference on*, pages 1–5. IEEE, 2016.

[19] KASV Rathnayake, WKIL Wanniarachchi, and WHKP Nanavakkara. Human computer interaction system for impaired people by using kinect motion sensor: Voice and gesture integrated smart home. In *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pages 531–536. IEEE, 2018.

[20] AMSB Adikari, NGC Ganegoda, and WKIL Wanniarachchi. Non-contact human body parameter measurement based on kinect sensor. *IOSR Journal of Computer Engineering*, 19(3):80–85, 2017.

[21] Jih-Gau Juang and Chang-Yen Yang. Document delivery robot based on image processing and fuzzy control. *Transactions of the Canadian Society for Mechanical Engineering*, 40(5):677–692, 2016.

[22] S.I.A.P. Diddeniya, W.K.I.L. Wanniarachchi, P.R.S. De Silva, and Ganegoda N.C. A study on mapping, localization and autonomous navigation with irobot create 2 on ros indigo. In *International Conference on Multidisciplinary Approaches 2018*, volume 5, page 5. FGS-USJP, 2018.

[23] Mihai Jalobeanu, Greg Shirakyan, Gershon Parent, Harsha Kikkeri, Brian Peasley, and Ashley Feniello. Reliable kinect-based navigation in large indoor environments. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 495–502. IEEE, 2015.

[24] Diogo Santos Ortiz Correa, Diego Fernando Sciotti, Marcos Gomes Prado, Daniel Oliva Sales, Denis Fernando Wolf, and Fernando Santos Osorio. Mobile robots navigation in indoor environments using kinect sensor. In *2012 Second Brazilian Conference on Critical Embedded Systems*, pages 36–41. IEEE, 2012.

[25] H Durrant Whyte. Simultaneous localisation and mapping (slam): Part i the essential algorithms. *Robotics and Automation Magazine*, 2006.

[26] Kimitoshi Yamazaki, Ryohei Ueda, Shunichi Nozawa, Mitsuharu Kojima, Kei Okada, Kiyoshi Matsumoto, Masaru Ishikawa, Isao Shimoyama, and Masayuki Inaba. Home-assistant robot for an aging society. *Proceedings of the IEEE*, 100(8):2429–2441, 2012.

[27] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, Japan, 2009.

[28] John J Leonard and Hugh F Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Transactions on robotics and Automation*, 7(3):376–382, 1991.

[29] Dieter Fox. Kld-sampling: Adaptive particle filters and mobile robot localization. *Advances in Neural Information Processing Systems (NIPS)*, 2002.

[30] Mike Schuster. Speech recognition for mobile devices at google. In *Pacific Rim International Conference on Artificial Intelligence*, pages 8–10. Springer, 2010.

[31] Leonard Richardson and Sam Ruby. *RESTful web services*. " O'Reilly Media, Inc.", 2008.