Integrating iRobot-Arduino-Kinect Systems: An Educational Voice and Keyboard Operated Wireless Mobile Robot

S. I. A. P. Diddeniya[1], K. A. S. V. Rathnayake[1], H. N. Gunasinghe[2], W. K. I. L. Wanniarachchi[1*]
[1]Department of Physics, Faculty of Applied Sciences, University of Sri Jayewardenepura, 10250 Sri Lanka
[2]Department of Mathematics and Computer Science, The Open University of Sri Lanka, 10250 Sri Lanka
Corresponding Email: iwanni@sjp.ac.lk

Abstract: This paper presents use of iRobot Create Programmable Robot for real time wireless controlling via speech and keyboard commands in a domestic environment. Commercial hardware such as iRobot create 2 programmable robot, Kinect V2, Arduino Uno and NRF24L01 radio frequency module were used, based on open-source code in the implementation. At the receiver end iRobot programmable robot was used as the mobile robot. At the transmission end, a PC based system was used. The keyboard controlling system is designed to control the directions of the moving robot. In the voice controlling unit, voice commands were recognized using the inbuilt speech recognition engine of the Kinect sensor. We provide comprehensive analysis of the performance of the developed robot system in an indoor environment. The results show that 89% of accuracy for untrained non-native (Sri Lankan) English speakers in British accent.

Keywords: Microsoft Kinect V2, iRobot Create, Speech Recognition, Voice Control, Arduino, Wireless Robot

I. Introduction

Researchers are continuously trying to develop robots that are self-sustaining and more intelligent to use in various applications. Robots do different tasks in many fields. According to the application they are involved in, they can be categorized into eight different categorizations such as Industrial robots, Domestic or household robots, Medical robots, Service robots, Military robots, Entertainment robots, Space robots and Hobby and competition robots. This proposed system uses a service robot based on iRobot Create Programmable robot which can move autonomously from one place to another according to the given voice or keyboard commands.

In the recent years, speech recognition is earning very much interest of people and the developments based on speech recognition has furnished more natural ways of communicating with machines, creating more easier and user-friendly environment. The system allows the user to communicate with the robot naturally by using predefined spoken words. So this proposed programmed robot recognizes voice commands and travel from place to place and achieves the goal without colliding or getting lost.

II. Related Work

There has been a massive progress in the area of autonomous robot navigation that demonstrates different capabilities for different purposes. This section will provide brief descriptions of various robot navigation systems that have been developed in

different time frames. Siliveru Ramesh *et al*. proposed a brain-computer interface system for a mind controlled robot. The system uses Mindwave Headset which is provided by Neurosky Technologies and those signals will be transferred by using Bluetooth and uses LabVIEW design software to analyze brain waves [1].

Muhammed Zahit *et al.* proposed a robot controlling system with voice commands that takes speech commands to the computer by using a microphone and the features are extracted by using Mel Frequency Cepstral Coefficients algorithms. The speech recognition part of the system was implemented with the help of Artificial Neural Networks and recognized command is delivered to the robot after it is being converted to the format which robot can identify [2]. The proposed system for fire extinguishing robot by M. Pednekar *et al.* used visual studio programs for speech recognition to control the robot [3]. Some proposed systems used android platform for speech recognition and accordingly control the robot [4].

Jayesh Chopade *et al.* proposed a system to control an imagined spy robot by voice and computer keyboard commands. The visual feedback is taken from the camera which is composed of the system used to provide the control to the robot more accurately. The proposed system used ZigBee technology and it is used to communicate with the voice recognition module which contains HM2007 IC. HM2007 IC recognizes up to 20 words and the length of that each word is about 1.92(s). So it narrows down the amount of vocabulary that speaker use. Some systems used MATLAB software for speech recognition process [5] [6]. As MATLAB is not open source software, it rises up the cost of the system. David Be et al, propose a system that can wirelessly control a LEGO NXT robot using voice commands. They have used voice recognition module which performs the task of speech recognition, through "CSLU TOOLKIT" and used Bluetooth communication module [7].

Studies in [8] [9] [10] contain some examples of a Kinect based systems and methods to navigate a robot by using Kinect gestures and voice. One system uses a humanoid robot which drives servo motors and the other one uses Bipedal Robot which has six servos at six joints providing desired walking movement. The proposed work in references [11], [12], [13], [14], and [15] systems were based on gestures that use Android applications and image processing where user can control the robot by using voice commands.

In this work, iRobot-Microsoft Kinect combination is used to build low cost wireless robot navigation system in a domestic environment. iRobot create programmable robot was used as the mobile platform. The movement of the robot can control via

voice and keyboard commands remotely. Microsoft Kinect voice recognition capabilities were used to identify the predefined voice commands. The proposed system can be used in hazard and unsafe areas where people cannot reach and can be used as a butler for disable people and older people.

III. Implementation

The proposed system has two main units including control unit and robot system unit. In order to implement the robot controlling system, Kinect version 2.0 was used with Visual C# programming. iRobot create 2 programmable robot was used as the main robotic component with Arduino support and it was reprogrammed to wireless control. The detail explanation of iRobot-Arduino interfacing and communication is given in later parts of the paper.

A. Control Unit

The controlling unit of the system regulates navigation of the robotic structure according to the user inputs provided either as voice commands or keyboard instructions. Microsoft Kinect API was used to construct the speech module which drives our audio system. The control unit consists of following four main parts.

1) Keyboard Functions:

To control the motion of navigating robot, we implemented keyboard controlling mechanism as shown in figure 1 considering, moving forward in a straight line and turning into four main directions, with respect to the current position. Mainly continuous pressing of key moves the robot and releasing results in termination. Directional keys turn robotic wheels in the relevant direction as given in table 1.

2) Voice Controlling:

We built a robot controlling system for voice commands using Microsoft Kinect sensor version 2.0 as the voice capturing device which is shown in figure 1. Its microphone array component was used in this proposed work. Voice signals captured by the microphone array are noise filtered and directional, hence gives a higher accuracy than a single microphone. Captured signals are further processed in a Kinect's inbuilt preliminary processor. The Kinect audio signals are then used for speech recognition using Kinect speech module.

3) Speech Module:

In the Speech module section, we have used Microsoft Speech Recognition Engine (Speech recognizer) and Microsoft Speech Recognition Grammar modules. We used British Language (EN-GB) and United States (EN-US) language accent packs for testing purposes. Voice commands were predefined for the control of robot navigation as stated in table 2.

4) Speech Recognition:

Microsoft provides a speech recognition engine that matches vocal inputs with the words and phrases, defined by grammar rules of speech recognition. Among simple and complex rules, simple rules recognize small set of words and phrases. A more complex grammar can recognize and organize semantic content from a variety of user accents. A speech recognition grammar

also defines a set of grammar specific properties, such as locale, semantic format, and mode.

Microsoft speech recognition engine handles the part of speech recognition with a confidence value which lies between 0 and 1. A given confidence threshold value is compared with the confidence value which depicts the certainty of the selection. If that confidence value is greater than or equal to the given confidence threshold value, the command will be executed [16].

B. Wireless Mobile iRobot System

As shown in figure 2 the robot system unit consists with an iRobot, which is programmable equipment that has two rotatable wheels. Additionally, a programmed Arduino Uno board with an NRF24l01 radio frequency module is mounted on the iRobot in order to make the robot wireless. Having a light weight structure of the mobile platform speed, power consumption and durability can be maintained.

1) iRobot:

iRobot, comes with an inbuilt program to navigate and operate as an automatic cleaning agent, to be able to operate wired or wireless, however can be reprogrammed for various purposes. Arduino boards, computers and raspberry pi can serially interface with iRobot for communication. Predefined op-codes ease the reprogram ability for both primary and derived operations on the robot. There are five opcode command categories for getting started, define operating mode, cleaning, control actuators, and input. Drive command (op-code ID: 137) which is a command to control actuators, defines using 4 bytes as Drive [velocity high byte] [velocity low byte] [radius high byte] [radius low byte]. In addition, input commands can be used to get data such as distance travelled, angle, voltage, current, etc. from the robot [17].

2) Interfacing Arduino to iRobot:

To make the robot navigate wirelessly by implementing a communication link between controlling unit and the robot, we used a NRF24L01 radio frequency modules (RFM). The RFM is transceiver that has 2.4 GHz frequency with a long radial length of 10 m without an antenna, that gives a fast communication and a longer travel length to the robot. As it uses radio frequencies, the cost is reduced.

Communication between iRobot and Arduino UNO was implemented using Arduino built in Software Serial library. Here, we assigned pin no. 10 as RX pin and pin no. 7 as TX pin of the Arduino board. The iRobot's TX pin and RX pin was connected to the Arduino Uno pin number 10 and 7 respectively. Baud rate pin of the iRobot was connected to pin 3 in Arduino UNO board. Additionally, power required for the Arduino board was obtained by the iRobot. Hence we connected V_{pwr} pin and GND pin of iRobot through 5V LM7805 regulator circuit to arduino board's V_{in} pin and GND pin [18].

TABLE I. Keyboard event handling

Event Keys	Up / W	Down / S	Left / A	Right / D	Enter	Space
Key Press	Move forward	Reverse	Turn left	Turn right	Power On	Shut Down
Key Release	Stop the given command for robot				-	-

TABLE II. Voice Commands

Command	Description
Forward	Move with 100 mms^{-1} speed of 0 mm radius (straight)
Backward / Back	Move with -100 mms^{-1} speed and 0 mm radius (straight)
Turn left	Rotate counter clock wise with 100 mms^{-1} speed of 1 mm radius
Turn right	rotate clock wise with 100 mms^{-1} speed of -1 mm radius
Stop	Reduce speed gradually until speed reach to 0 mms^{-1}
Speed up	speed = current speed + 1 mms^{-1} until 500 mms^{-1}
Slow	speed = current speed -1 mms^{-1} until 0 mms^{-1}
Power off	Decrease speed until 0 mms^{-1} and shut down the iRobot
Robot on / Start	iRobot start

Fig. 1. Robot controlling system using keyboard commands and voice commands

3) Controlling iRobot through Arduino:

An Arduino Uno board was programmed to transmit signals from computer side to mobile robot. Speech module/keyboard sends the recognized command to the Arduino Uno via serial (RS232) communication. Accordingly the selected command is executed. This system can be changed at any time by defining new commands. In our proposed work we have used tentative four outputs which are move forward, move backward, turn left and turn right. If we need to drive iRobot straightly with 100 mms^{-1}, then we need to send following commands to iRobot from the arduino program.

```
int velocity = 100; //set the speed in mm/s
int radius = 0;     //set the radius to 0 mm
Serial.write (137); // opcode of drive function
Serial.write (highByte (velocity));
Serial.write (lowByte (velocity));
Serial.write (highByte (radius));
Serial.write (lowByte (radius));
```

Following arduino code was used to drive in reverse with velocity of 200 mms^{-1} while turning left at a radius of 500 mm.

```
int velocity = -200;
int radius = 500;
Serial.write (137);
Serial.write (highByte (velocity));
Serial.write (lowByte (velocity));
Serial.write (highByte (radius));
Serial.write (lowByte (radius));
```

IV. Experiment and Performance

An experiment was conducted and data analysis was made to test the proposed work in order to get an idea of the system accuracy and how relevant it is to the practical world. Accordingly the developed system was tested with Kinect's EN-GB and EN-US language accent packs. Our tested environment was isolated of dimension 12 m x 8 m x 6 m.

Nine voice commands from four individuals consist of two males and two females were used for the experiment in order to increase the system's generalization. The individuals' sound level of the voice commands was between 65 dB - 80 dB as sense by Kinect. The performance distance between user and the Kinect sensor (Microphone Array) was a constant of 0.5 m apart. Confidence value was set to 0.3. The user was asked to pronounce each of the predefined voice commands. We identified the accuracy of the recognition from the iRobot platform movements according to one of two possible outcomes; executing exactly as the given command, executing any other command. The experiment was repeated ten times per command for each person to increase the accuracy of the experiment. Figure 3 and 4 show the variation of recognition accuracy for each person per each command for two chosen English accents. In the graphs shown in figure 3 and 4, blue, red, green and purple bars represent female user 1, male user 1, female user 2 and male user 2 respectively. Solid blue line in each graph shows average recognition as a percentage. It is illustrated that in EN-GB training, the accuracy of every command was more than 50% regardless the gender of user. Overall, 89% and 61% of accuracy was shown for the set of voice commands, for EN-GB
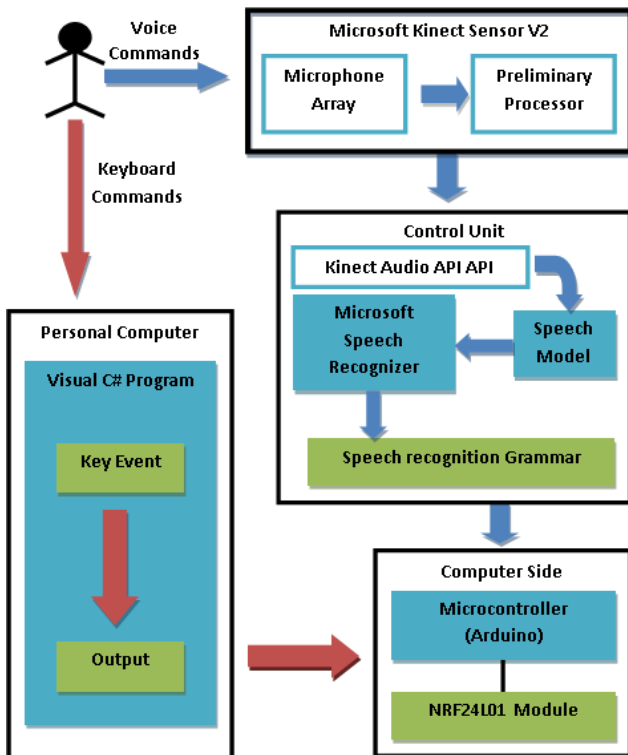
and EN-US respectively. The accuracy of EN-GB is greater than 85% for all users whereas that was between 50%-69% for EN-US language system.
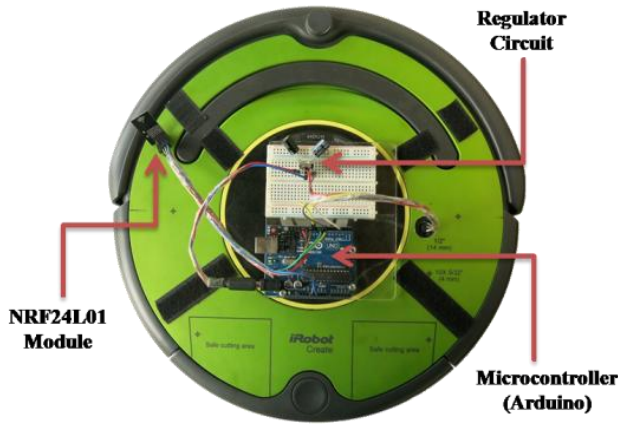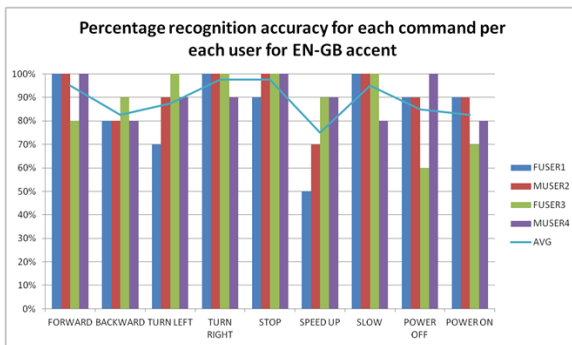

Fig. 2. Wireless Mobile iRobot System

Fig. 3. Percentage recognition accuracy for each command per each user for EN-GB
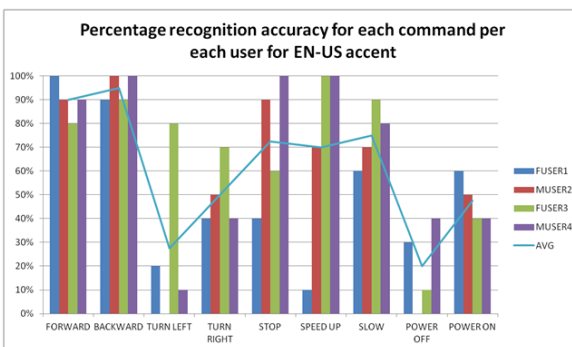

Fig. 4. Percentage recognition accuracy for each command per each user for EN-US

V. Discussion

The voice recognition accuracy of EN-GB is greater than that of EN-US for randomly selected untrained non-native English speakers (Sri Lankan users), implies that these users' accent is more similar to EN-GB than EN-US. The result depends on the confidence value which defines the similarity of a vocal to that of a British or American. If the confidence value is high, then the commander's speech accent must be closer to native English pronunciation. As we set a low confidence value, a wide range of accents become valid to the recognizer.

Choosing meaningful voice commands is important to human robot interactions. Due to the Kinect's powerful voice recognizer, one can extend voice instructions through words, phrases and sentences by using multiple voice commands for a single functionality. It allows a range of words to control the robot hence enhancing the usability of the robot.

Keyboard controlling of the system, limits its capacity to basic controlling aspects except speed controlling. Execution time of keyboard controlling is lesser than that of voice controlling.

In the voice controlling process, we used speed up, slow, stop commands. In speed up command, robot velocity is increased step by step until robot velocity reaches its maximum speed 500 mms^{-1} from current velocity. If robot's current velocity is zero (robot in stop mode) or positive (robot in forward mode) then its velocity is increased to forward direction. In contrast, if the current velocity is negative (robot in reverse mode) then the velocity increases to backward direction. Similarly in slow command, the robot velocity decreases step by step until it becomes zero in both forward and reverse modes. In addition, steps of slow command are equal to steps of stop command. The pseudo code of speed up command is given bellow.

```
int currentVelocity = velocity regarding to previous command;
int currentRadius = 0;
// because robot can speed up only forward and backward
if (Robot is moving forward or robot is stop)
// currentVelocity >= 0 (positive)
{
        while (currentVelocity < 501)
        {
        Serial.write (137);
        Serial.write (highByte (currentVelocity));
        Serial.write (lowByte (currentVelocity));
        Serial.write (highByte (currentRadius));
        Serial.write (lowByte (currentRadius));
        currentVelocity = currentVelocity + 1;
        }
}
else (Robot is moving to backward)
// currentVelocity < 0 (negative)
{
        while (currentVelocity > -501)
        {
        Serial.write (137);
        Serial.write (highByte (currentVelocity));
        Serial.write (lowByte (currentVelocity));
        Serial.write (highByte (currentRadius));
        Serial.write (lowByte (currentRadius));
        currentVelocity = currentVelocity - 1;
        }
}
```

VI. Conclusion

In this work, we used keyboard commands and voice commands to control the iRobot mobile platform while enabling wireless communication between the two units. The detail explanation of interfacing iRobot and Arduino Uno is given above. We conducted a laboratory scale experiment on voice controlling using Kinect V2 with EN-US and EN-GB language packages. Recognition accuracy for EN-GB was 89% and 61% for EN-US. Hence it can be concluded that Sri Lankan English accent is

closer to EN-GB than to EN-US. Further, in our system the maximum controllable distance between wireless robot unit and the control unit was 10 m at which when we used NRF24L01 modules without an antenna.

Acknowledgement

References

i. S. Ramesh, M. G. Krishna, and M. Nakirekanti, "Article: Brain computer interface system for mind controlled robot using bluetooth," International Journal of Computer Applications, vol. 104, no. 15, pp. 20–23, October 2014, full text available..

ii. M. Zahit, "Robot control with voice command," Senior Project, Yıldız Technical University, Department of Computer Engineering, Istanbul, 2008.

iii. M. Pednekar, J. Amanna, J. John, A. Singh, and S. Prajapati, "Voice operated intelligent fire extinguishing vehicle," in 2015 International Conference on Technologies for Sustainable Development (ICTSD), Feb 2015, pp. 1–4.

iv. M. Narayana, A. Alishety, and H. Chapala, "Voice control robot using android application," International Journal of Engineering Innovation & Research, vol. 4, no. 2, pp. 332–337, 2015.

v. J. Chopade, D. Barmade, and S. Tonde, "Control of spy robot by voice and computer commands," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, no. 4, pp. 1835– 1837, April 2013.

vi. S. D. Gosavi, U. P. Khot, and S. Shah, "Speech recognition for robotic control," International Journal of Engineering Research and Applications, vol. 3, no. 5, pp. 408–413, Sep-Oct 2013.

vii. D. Be, M. Garcia, C. Gonzalez, C. Miranda, M. Escalante, and S. Gonzalez, "Wireless control lego nxt robot using voice commands," International Journal on Computer Science and Engineering, vol. 3, no. 8, pp. 2926–2934, August 2011.

viii. S. N. Mahmood, N. A. Siddiqui, M. T. Qadri, S. M. Ali, and M. Awais, "Humanoid robot control by body gestures and speech," in First International Young Engineers Convention. Lahore: FEIICIYEC, April 2014.

ix. N. James and R. M. Jose, "Gesture controlled bipedal walking robot using kinect," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 4, no. 7, pp. 6536– 6541, July 2015.

x. Z. Lei, Z. H. Gan, M. Jiang, and K. Dong, "Artificial robot navigation based on gesture and speech recognition," in Proceedings 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), Oct 2014, pp. 323–327.

xi. A. Ganihar, S. Joshi, G. Rahul, R. Hongal, and U. Mudenagudi, "Android based wireless gesture controlled robot," in 2014 International Conference on Advances in Electronics Computers and Communications, Oct 2014, pp. 1–4.

xii. P. Chanda, P. K. Mukherjee, S. Modak, and A. Nath, "Gesture controlled robot using arduino and android," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 6, no. 6, pp. 227–234, June 2016.

xiii. H. K. Kaura, V. Honrao, S. Patil, and P. Shetty, "Gesture controlled robot using image processing," International Journal of Advanced Research in Artificial Intelligence, vol. 2, no. 5, pp. 69–77, 2013.

xiv. J. Sheetal, K. Payal, J. Bhagyashree, and T. Azhar, "Hand gestures based car control system," International Journal of Advancement in Engineering Technology, Management & Applied Science, vol. 2, no. 1, pp. 96–101, January 2015.

xv. S. A. Urabinahatti, M. R. Hudagi, Sharankumar, and R. Gujanatti, "Robot navigation control using hand gestures," International Journal of Modern Trends in Engineering and Research, vol. 2, no. 6, pp. 320– 327, June 2015.

xvi. "About speech recognition grammars (microsoft.speech)," https://msdn.microsoft.com/en-us/library/hh378438(v=office.14).aspx, accessed: 26 September 2016.

xvii. iRobot Create 2 Open Interface (OI) Specification based on the iRobot Roomba 600, iRobot Corporation, http://www.irobot.com/, April 2015.