

# Mobile Robot Voice Recognition in Control Movements

Zakariyya Hassan Abdullahi, Nuhu Alhaji Muhammad, Jazuli Sanusi Kazaure, and Amuda F.A.

**Abstract**—Speech Recognitions a prominent technology for Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI). The increase in the use of robots and automation has significantly attracted the attention of both academic research and industrial applications. In addition to facilitating the daily work the use of robots and automation has helped the productivity and reduces wastage. Although there are many ways that were developed to communicate with the robot, but the ability to communicate verbally can provide a new communication approach. The main objective of this paper is to develop a system that is able to recognize the voice of a consumer and to control the robot movements with verbal instructions. This paper has chosen to use a novel method known as waterfall model. This approach gives a rigorous viewpoint and easy implementation of voice recognition system. The application of voice recognition plays an important role in efforts to convert the oral instructions referred to by the words that can be identified by the System translated into Arrick robot command.

**Keywords**—Intelligence, Human-Robot Interaction (HRI), Human-Computer Interaction (HCI), waterfall model

## I. INTRODUCTION

The term of Social interaction and intelligence is important in our daily life. To an Artificial intelligence and Robotics community it is one of the challenging areas in Human-Robot Interaction (HRI). Speech recognition technology is a great point to admit. The challenge and it is a prominent technology for Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI) for the future. Speech recognition is a technology. Where the systems understands the words (and its meaning) given through speech.

The increase in the use of robots and automation has significantly altered our perspective on it. In addition to facilitating the daily work, the use of robots and automation has helped the productivity and reduces wastage. One important aspect in the use of robots is the interaction between robots and humans. Many robots depend on the people to give detailed instructions in order to perform certain functions. There are many ways used to give instructions to a robot or a machine to perform a task. Among the methods

used in general are the use of physical devices such as keyboard, mouse and the use of sensors such as voice sensors; motion sensors, temperature sensors and so on. The increase in the smooth interaction between robots and humans can increase in the operation of a robot.

Although there are a lot of researches in improving the interaction between humans and robot but the smoothness of the interaction is still in need of repair in the future. Robots are expected to interact with men as a man talking to each other. Although it has not yet reached such a level but now the robot can receive commands from users verbally.

There are various systems and robots that have applied the method of voice recognition but it is limited. One of the smartest robots which are Honda's ASIMO has made the voice recognition system as described. The robot ASIMO can recognize when her name was called and facing the sound source see the faces of those speaking and answer, then to recognize unusual noises that occur suddenly like a falling object or a collision and face in that direction.

## II. RELATED WORK

### A. Background

The recent technology in robotics attracted the attention of both academic researchers and industry domains [1], [2], [3], [4], [5]. Most of the robots are primarily designed to assist humans performing some task for instance Medical robots used in various operation; removing of plaque from patient's arteries etc. [5]. Therefore, it is imperative to build a common interface such as human-robot interaction (HRI). Application of natural speech via speech recognition technology is possible with HRI system. However, there is yet some constrain in achieving the perfect HRI system because challenges involved in integrating speech recognition systems with the dialogue managers of robots [6]. Service robots helping aged people, in particular becomes the focus of today's research [7], [8], [9] because of the systematic increase in the dominant old population as well as the increase in the costs of the ageing people care. Some of these service robots have been developed as outcomes of some research works carried out in the developed nations [7], [10]. Another work[11]implemented the voice recognition systems in the ATmega162 microcontroller. Words (instructions in the voice) in a sign language used to control the movement of a mobile robot. Mobile robot will move according to the signal command language, there are five commands,

Zakariyya Hassan Abdullahi, Jazuli Sanusi Kazaure and Amuda F.A., Hussaini Adamu Federal Polytechnic Kazaure, Nigeria  
Email id: zakariyyahassan41@yahoo.com, jazsak@gmail.com, babanla\_11@yahoo.com

Nuhu Alhaji Muhammad, Kano State Polytechnic, Nigeria, Email id: engineer\_nuhu@yahoo.com

Indonesian language used to control the movement of a mobile robot. Voice commands used are "advanced", "backward", "left", "right", and "stop" to control the mobile robot to move forward, move backward, turn left, turn right, and stop moving.

### B. Voice Recognition

The voice recognition application is an application used to translate the words referred to the text. This application is used to control computers by developing software that is able to operate by using verbal commands from the user. The analysis of speech or voice recognition has been investigated as a speech patterns between humans and machines. Progress in the understanding of speech rules coding, transmission, and voice recognition have been instituted from the beginning of this century, According to survey of the state of art in human language technology Cambridge University isbn-0.521 2006, Speech recognition is the process of converting an acoustic signal captured by microphone or a telephone to a set of words. There two important part in Speech Recognition;

i) Recognize the series of sound and ii) Identified the word from the sound. This recognition technique depends also on many parameters - Speaking Mode, Speaking Style, Speaker Enrolments, Size of the Vocabulary Language; Model Perplexity; Transducer etc. There are two types of Speak Mode for speech recognition system- one word at a time (isolated-word speech) and continuous speech. Depending on the speaker enrolment, the speech recognition system can also divide - Speaker dependent and Speaker independent system. In Speaker dependent systems user need to be train the systems before using them, on the other hand Speaker independent system can identify any speaker's speech.

Vocabulary size and the language model also important factors in a Speech recognition system. Language model or artificial grammars are used to confine word combination in a series of word or sound. The size of the vocabulary should be in a suitable number, Large numbers of vocabularies or many similar sounding words make recognition difficult for the system.

### C. Robot Arrick

Arrick Robots will be used as a robot controlled by voice. According Arrick [12] there are two sets of robots that can be used for development, robots that can be programmed and robots that cannot be programmed Examples of which cannot be programmed robot is a robot Soccer Jr., Hyper Line Tracker Robot and others Robots that cannot be programmed, in general cannot be modified by the user and only run on a command that has been set by the manufacturer. Examples of robot that can be programmed are a robot Arrick that going to be used. In this project Programmable robots have the advantage to change the program and add some futures works that can be controlled. Arrick Robots can be programmed to follow the instructions given by voice. For the purpose of programming the robot Arrick manufacturers have been producing IDE known as the BASIC Stamp Editor (version 2.2.5) using Visual Basic programming language. The program developed with BASIC Stamp Editor is used to get

voice command has been processed to control the robot movement, Processing voice commands or voice recognition applications will be controlled by the Microsoft Speech SDK programmed with the language of Visual Basic (VB).

### D. Microsoft Speech SDK

Microsoft Speech SDK, use to assist in the development of voice recognition applications system, According to Microsoft [13] Microsoft Speech SDK is a voice recognition engine developed by Microsoft for voice recognition applications can be integrated in the development of a system. Voice recognition engine has been used to detect sound from the microphone and turn them into words; it is also used to list the words in a library or storage that can be detected by voice recognition applications.

Microsoft Speech SDK implemented, in the process of voice criticizes that give Instructions to the robot, Earphones as a microphone will use to get instructions from the user. After that the voice commands will be processed using the voice recognition application developed using the Microsoft Speech SDK, after processing voice commands are translated to BASIC commands to be sent to the robot.

The robot will react according to the instructions given. The main purpose of using Microsoft Speech SDK its enables changes in the form of voice commands that can be interpreted to robot commands.

## III. PROPOSED METHODOLOGY

This project has chosen to use a development methodology known as waterfall model because these models are easy to use and good for control. This approach has six major phases as shown in Fig. 1.

Waterfall model (Waterfall Capital) has been chosen to develop this project because this model can classify the level or phase of development carefully so that the smoothness of the project can be improved.

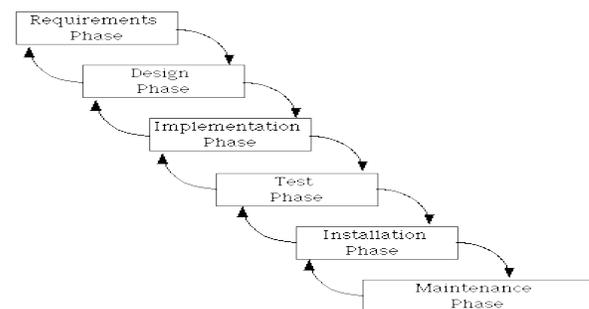


Fig. 1 Water fall methodology

### A. Software Requirements

This phase is the first stage in the development of this system. This phase aims to identify the scope and boundaries of the problem and plan goals strategies for system simulation. Initiated by the project development needs research projects will be conducted to find existing software

can be used for recognize the voice of the user, the voice that has been identified should be translated into robot commands used in order to act like a question, academic exercises similar to this software will be compared in order to identify strengths and weaknesses that affect the quality of application software.

*B. Design Phase*

The main activity in this phase is the development design for various components of the application architecture and interface architecture, this phase is to determine specifically how the application will work. Figure structure of the application will be built together with the flow chart for the module structure of the application; the voice recognition module, the module controls the robots and computer communications module.

*C. Implementation Phase*

This phase serves to develop and provide applications to operate. The main activities in this implementation phase is programming that will realize all the planning activities that have been made. Microsoft Speech SDK is used in the development of speech recognition. Basic Stamp Editor is used to build the code and download on the micro-robot controllers. Programming language Visual Basic (VB) is used to develop the application interface and control functions of the relationship between robots and computers.

*D. Testing Phase*

The main purpose of this phase is to test whether the software has been developed capable of operating as planned or not. Software errors can be avoided will be identified and the correct program in the implementation phase. For example during the application be passed and serial port that connects to the robot on a computer is not set correctly will display a message telling specified serial port cannot be used.

*E. Software Maintenance Phase*

In this phase, maintenance activities are the main activities carried out among the maintenance done, including corrections, adjustments and refinements. The image obtained from the user will be considered to improve the completeness of this application. This phase is important to ensure that the software runs smoothly without any problems.

IV. EXPERIMENTAL PROCEDURE

Voice recognition module, as shown in Fig. 2 using the Microsoft Speech SDK as the voice recognition engine, when the application interface is switched on a set of words to be recognized is sent on the voice recognition engine storage and identification system will be activated.

Communication between robot and Computer Company is run using the wireless with Bluetooth. Fig. 2 shows the communication module, simply robots and computers. Bluetooth will be installed in the robot and the computer also. After the robot is turned on, Bluetooth on the robot will also be turned on. After that the robot will be detected on the computer and connected via the com port. At the interface the

com port that connects between the robot and computer must be selected, and connected to the system.

Robot communication module and computer will start after the voice recognition module identify and send word to the system. By using the word returned, landing in one of the characters to be sent through the serial port to be determined.

Robot control module commenced after the micro controller of the BASIC Stamp II robot to receive commands from the computer on. Instructions on the robot will arrive in the form of characters. Each character is used to control the movement of different robots activities.

Table I, explains the function of each character in robot control. When a character is received by the microcontroller on the robot, the microcontroller will examine characters that are sent and respond accordingly.

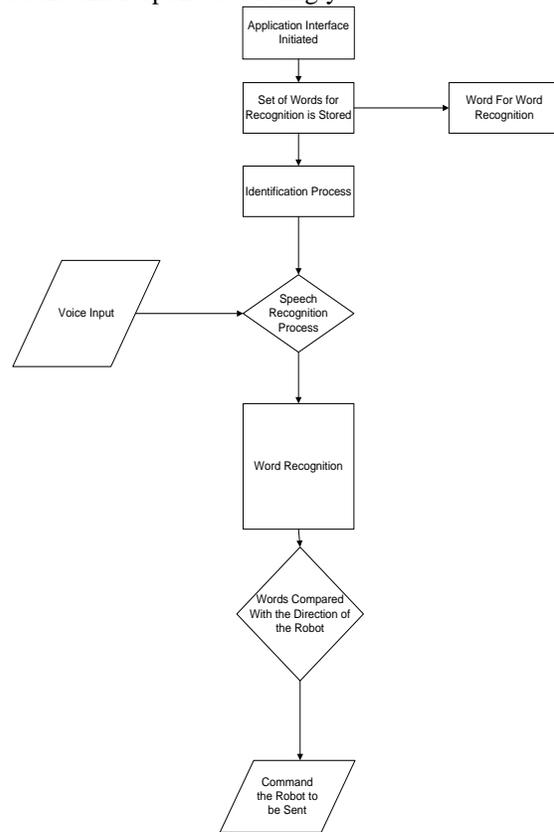


Fig. 2 Voice recognition module

TABLE I  
FUNCTION OF ROBOT CONTROL CHARACTERS

Script	Robot Character Response
A	Turning a green LED
B	Turning off green LED
C	Switch the red LED
D	Off the red LED.
E	Sounding beep once
F	Sounding f beep twice
G	Turn to the left
H	Turn to the right moment

Script	Robot Character Response
I	Move straight
J	Move straight
K	Move forward at low speeds
L	Moving forward at high speed
M	Move back on medium speed
N	Mobile back at low speeds
O,P	Move back at high speed

V.SYSTEM IMPLEMENTATION AND TESTING

This section includes activities of implementation and testing of applications, the study of this phase involves coding and testing. Encoding refers to the translational logic of application design, program code using appropriate programming language. Applications tested repeatedly by using various methods of testing to find and fix any errors that may arise. This will describe the application of flow testing applications and syntax. This section discusses the methods and programming languages used in the application development process. Testing methods were used in the application stage.

A. Coding

The language used in the project Voice Recognition Applications Limited as the Robot Movement Control is a Visual Basic (VB) and Basic stamp editor, interface part voice recognition and transmission of information developed using VB. Meanwhile, the receipt of information or instruction, direction and control of robotic processing developed with Visual Basic. Coding part will provide information in accordance with the modules. In the application of voice recognition module, communication module and the computer and robot control modules shows in Fig. 3.

```
PublicSub ProcessCommand(ByVal
command AsString)
If _serialPort.IsOpen Then
serialPort.WriteLine(command)
```

Fig. 3 Code for sending command in Robot

B. Application Interface

The application interface can be divided into two important parts of the robot and the relationship with information. About covering parts of information texts, that describes the voice commands that can be given to the robot. This section does not have any control on the application. Part Connected to Robots are crucial in the application interface. Most of the control system can be modified in this section.

Fig. 4 shows the information in the application interface.

Voice input control section has two parts, an indicator that shows the height of the input voice that works when there is no voice input and the second is the height of the control

input voice using a voice track bar to the height can be changed according to circumstances. There are indicators that show a high level of voice as the track bar. Parts of the voice recognition are label that will indicate the operation being carried out by the process. After a word is detected, the word will be displayed in this section. Fig. 5 shows the word "Hello" is displayed on the interface after the application is detected by voice recognition applications.

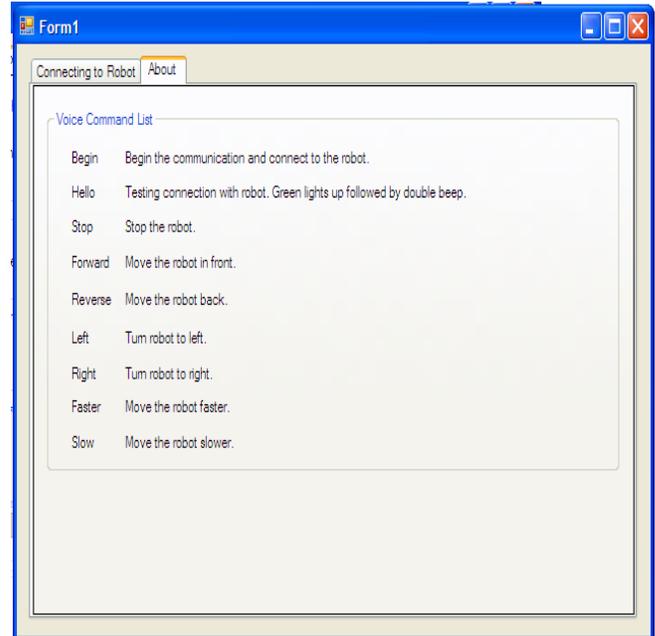


Fig. 4 Interface application divisions

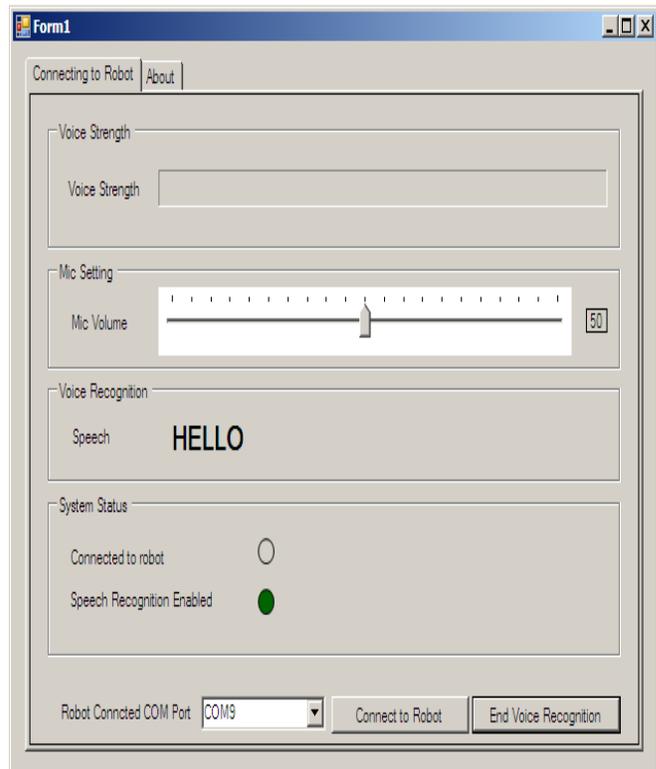


Fig. 5 The word "Hello" Trace and Displayed

Parts of the system status indicator will show the status of both the voice and the relationship with robots and computers. If the circle next to the label “Connected to Robots” is green, it will show the applications on the computer is connected to the robot and ready to send commands to the robot. If the circle next to the label "Speech Recognition Enabled" in green indicates that the voice recognition application is active and can take voice input for processing. No color in the circle indicates the application is not active. Fig. 5 shows the interface when the application on a computer connected to the robot and the active voice recognition applications.

The final part of the application interface is the communication and control buttons. This is the most important part of the application interface. This is because the control to start voice recognition system and communication with the computer carried out here. Dropdown will list all existing serial port on the system. The serial port where the robot is connected must be determined so that the communication between robots and computers running smoothly. After selecting the appropriate serial port, the "Connect to Robot" should be pressed to communicate with the robot. If communication is successful, the circle next to the label "Connected to robot" button will be green and will change the role to decide ties in the robot.

*C. Hardware and Installation Testing*

Several conditions must be followed for good installation of hardware parts; Application Project Voice Recognition in Restricted for Robot Movement Control will run smoothly.

- 1) Robot Arrick or ARobot must be installed according to instructions given. Some parts of the robot can be avoided such as whisker, which will not be installed because there is no use.
- 2) Bluetooth modem must be installed on the RS232 port on the robot and turned on so that communication can be initiated with a USB Bluetooth adapter on the computer. Fig. 6 shows the Bluetooth connection on board the robot controller.
- 3) Bluetooth modem on the robot and computer must be Bluetooth Adapter connected to serial port before it is ready for use by the system. Appendix D will show how the connection between the connected Bluetooth.

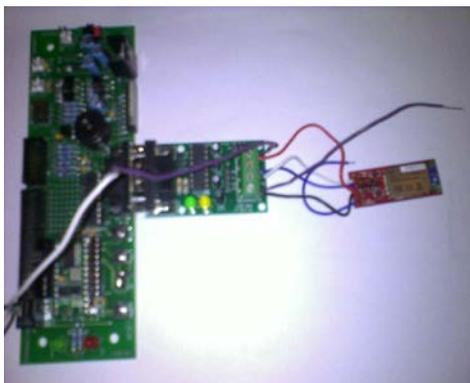


Fig. 6 A Bluetooth connections on Robot Controller Board

Application testing process conducted in parallel with the encoding process applications, especially to find syntax errors. Through testing, application weaknesses can be identified and application capabilities are unknown. Tests carried out to ensure it is free of errors, which can cause the application not work properly. Table 2 shows the times test for Voice Recognition Applications.

TABLE II  
TEST FOR VOICE RECOGNITION APPLICATIONS

Data	Word Recognition (%)
<i>Hello</i>	90
<i>Stop</i>	80
<i>Left</i>	75
<i>Right</i>	75

After testing the application, Mat lab software use to simulate the frequency signal for each command that has been tested. Are presented in Fig. 7, Fig. 8 and Fig. 9 respectively.

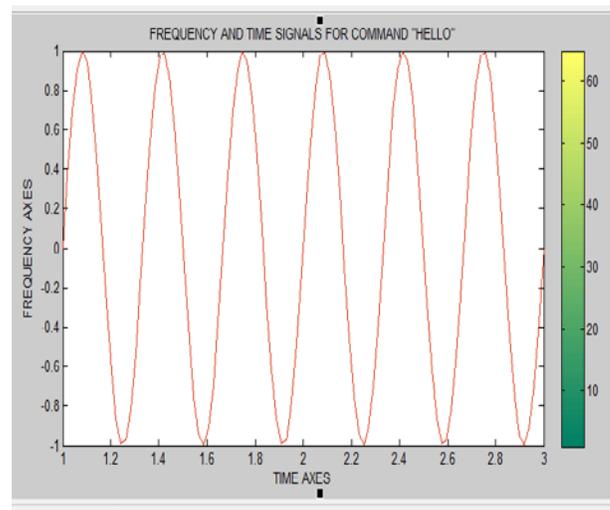


Fig. 7 Signal command “HELLO”

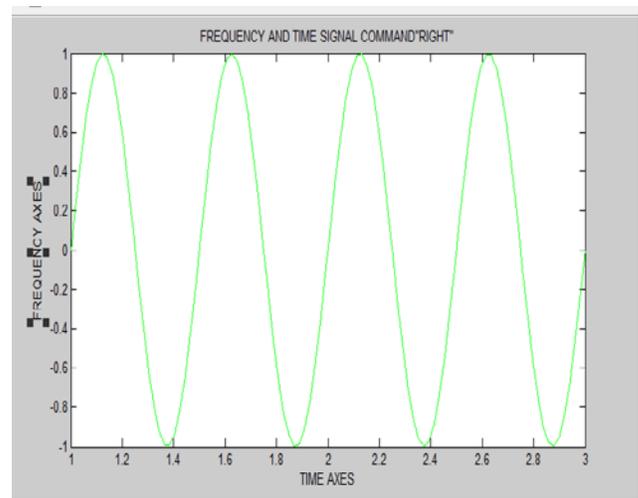


Fig. 8 Signal command “RIGHT”

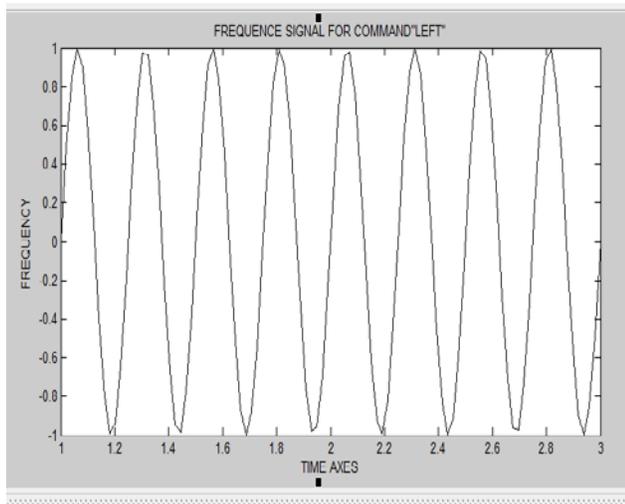


Fig. 9 Signal command "LEFT"

## VI. DISCUSSION

Overall objectives set in the early stages of this work have been achieved. Voice Recognition Application Project Limited as the Robot Movement Control was able to control the robot motion using the voice. The information provided in the study design can reflect the develop methods and functions contained in this application. The algorithm and flow chart are the key techniques, in making the description of the application design. Software and hardware specifications identified, these applications developed with the Microsoft Speech SDK 5.1 for voice recognition applications that play an important role. Special hardware provided in the development of this application as Arobot, Bluetooth modem, Bluetooth USB Adapter and a microphone. The accuracy of voice recognition applications,

Voice recognition applications have yet to achieve 100% detection rate. From Table 2, it can be concluded that the average rate or the accuracy of voice recognition is 80%. In addition, environmental noise can interfere with the voice recognition process run smoothly; the selected word is different syllables to improve the accuracy of voice recognition rates.

## VII. CONCLUSION

This study has achieved the objectives set at the beginning of Rating and it's attractive because it can interact with the robot, naturally. Although the developed application still has limitations and shortcomings, but these applications can be modified in the future for other uses such as wheelchair controlled by voice, entertainment robot which reacts with human voice and so forth. A number of advantages and constraints of the application were detected to enable the improvement of this application in future. Application development is expected to help students who wish to develop similar applications using another language.

## REFERENCES

- [1] R. Bischoff and V. Graefe, "Dependable multimodal communication and interaction with robotic assistants," in IEEE International Workshop on Robot and Human Interactive Communication, 2002.
- [2] T. Portele, S. Goronzy, M. Emele, A. Kellner, S. Torge, and J. Vrugt, "SmartKomHome - An advanced multi-modal interface to home entertainment," in Proceeding of European Conference on Speech Communication and Technology (EuroSpeech), Geneva, Switzerland, 2003.
- [3] I. Toptsis, A. Haasch, S. Huwel, J. Fritsch, and G. Fink, "Modality integration and dialog management for a robotic assistant," in Proceeding of European Conference on Speech Communication and Technology (EuroSpeech), 2005.
- [4] J. Ido, Y. Matsumoto, T. Ogasawara, and R. Nisimura, "Humanoid with interaction ability using vision and speech information," in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006.
- [5] C. Jayawardena and et. al, "Deployment of a service robot to help older people," in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, 2010, pp. 5990–5995.
- [6] M. Doostdar, S. Schiffer, and G. Lakemeyer, "A robust speech recognition system for service-robotics applications," in Proceedings of International RoboCup Symposium, 2008, pp. 1–12.
- [7] R. Reddy, "Robotics and intelligent systems in support of society," IEEE Transactions on Intelligent Systems, vol. 21, no. 3, pp. 24–31, 2006.
- [8] M. Kim, S. Kim, S. Park, M. Choi, M. Kim, and H. Gomaa, "Service robot for the elderly," IEEE Robotics and Automation Magazine, pp. 34–45, 2009.
- [9] C. Granata, M. Chetouani, A. Tapus, P. Bidaud, and V. Dupourque, "Voice and graphical -based interfaces for interaction with a robot dedicated to elderly and people with cognitive disorders," in Proceedings of international IEEE RO-MAN conference, September 2010, pp. 785–790.
- [10] B. Siciliano and O. Khatib, Springer Handbook of Robotics. New York: Springer, 2008.
- [11] T. Dhanny Wijaya, "Limited Speech Recognition for Controlling Movement of Mobile Robot Implemented on ATmega162 Microcontroller" Proceedings of the 1st Makassar International Conference on Electrical Engineering and Informatics 003CE, Hasanuddin University Makassar Indonesia, 2008.
- [12] R. Arrick, Robot Building for dummies Indianapolis, Indiana: Wiley Publishing, Inc. All rights reserved, 2003.
- [13] Microsoft. Of 2011, Microsoft Speech Application SDK (online) <http://msdn.microsoft.com/en-us/library/ms986944.aspx>, 2011.

About Author (s):



**Zakariyya Hassan Abdullahi**<sup>1</sup> became a Member of IEEE in 2003. He was born in Kano State, Nigeria in 1978. He received HND Electrical Engineering (2003) from Kaduna Polytechnic Nigerian and postgraduate diploma in Electrical Engineering (2008) from Bayero University, Kano (BUK), and an M.Sc. in System Engineering from University of East London (2012). Major field of studies is control engineering system.

He has work extensively in the laboratories as a senior 'TECHNOLOGIST' in Kano State Polytechnic and currently a lecturer II with Hussaini Adamu Federal Polytechnic Kazaure Nigerian. He published many articles including: Electronic communication system (with Hussaini Adamu Federal Polytechnic Kazaure Nigerian, December, 2012) His current research interest includes speech recognition in robotics and control systems.

Mr. Abdullahi is also a member of Nigerian Association of Technologist since 2005.