International Journal of Software Engineering and Its Applications
Vol. 8, No. 10 (2014), pp. 143-152
http://dx.doi.org/10.14257/ijseia.2014.8.10.13

The Framework of Navigation and Voice Recognition System of Robot Guidance for Supermarket

Satrio Halim and Widodo Budiharto

School of Computer Science, Bina Nusantara University – Jakarta, Indonesia
wbudiharto@binus.edu

Abstract

This paper presents the research on the development of framework for navigation and voice recognition system designed for robot in store/supermarket. The proposed system consists of a microcontroller and a voice recognition processor that can recognize many voice patterns. The robot is controlled through voice recognition system which allows customers to interact with the robot, in order the robot able to shows the desired location. The advantage of voice recognition system on a robot guidance are fast data input operation and easy to use. The compass is used for navigates the robot from start to goal position. The voice recognition system is trained in such a way that it is recognizing commands and the robot navigates based on the instruction through the voice commands. Based on the experiment, the system run very well and we evaluate and show the performance of this system.

Keywords: *robot guidance, voice recognition, navigation*

1. Introduction

The theme of Social interaction and behavior-based robotics is important and interesting to an Artificial intelligence and Robotics community. It is one of the challenging areas in Human-Robot Interaction (HRI) [1]. There is an advanced technology called Voice Recognition System that can be used and developed to make interaction between robots and humans. The characteristics of a human being are to have the ability to learn and evolve towards new patterns and stimuli. So here will be developed on a project about a robot that can understand and mimic the human ability. In this case the system will be developed into a robot guidance that will serve customers in a supermarket. The system currently is designed to understand human words.

Natural Language interface is now starting to appear in standard software application. This gives benefit to novices to easily interact with the standard software in HCI field. It also encourage expert in robotics to use Speech Recognition (SR) technology for the HRI. The way the system captures human words, so the word is accepted by a voice - responsive element in the microphone, the microphone change the variations of sound into electrical signals and voltage. Then the signals sampled and change into a digital bit stream. In this paper, we develop a system for robot guidance, which serves as an aid in finding the area of human goods available in store/supermarkets. Therefore writer developed a robot that is coupled with a voice recognition system that can show where the location of items to be searched by customers. Fezari *et.al.*, [2], propose a system consists of a microcontroller and a voice recognition processor that can recognize a limited number of voice patterns. The commands of autonomous robots are classified and are organized such that one voice recognition processor can distinguish robot commands under each directory. A voice

ISSN: 1738-9984 IJSEIA

command system is implemented with a microcontroller, a voice recognition processor RSC364 from Sensory and a set of radio frequency emitters- receivers. Unfortunately, the robot still not implemented for real life. In this research we propose a guidance robot that can be used in store/supermarket, in order the customer have a direction that guided by the robot.

2. Concept of Robot Guidance

2.1. Omniwheel Drive Systems

A holonomic or omni-directional robot is capable of driving in any direction. If we want to prescribe the robot's movements in the environment, we need to know how these variables relate to the primary variables we can control: The angular positions and velocities of the wheel shafts. Therefore, a kinematical model of the robot has to be developed. A global frame [x, y] represents the environment of the robot and the robot's location can be represented as (x, y, θ).

The global velocity of the robot can be written as $\dot{x}, \dot{y}, \dot{\theta}$. The center of this local frame coincides with the center of gravity of the robot. The three omni-wheel are located at an angle ai (i = 1, 2, 3) relative to the local frame as shown in Fig. 1. If we take the local axis xL as starting point and count degrees in the clockwise direction as positive, we have α1 = 0, α2 = 120 and α3 = 240°C [3]. The translational velocities of the wheels vi on the floor determine the global velocity of the robot in the environment $\dot{x}, \dot{y}, \dot{\theta}$ and vice versa. The translational velocity of wheel hub vi can be divided into a part due to pure translation of the robot and a part due to pure rotation of the robot:

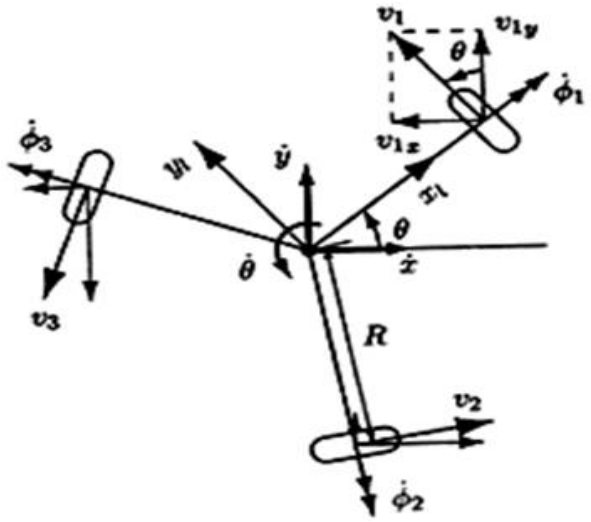


Figure 1. Kinematics Model of Omnidirectional Wheels

Motor driver is part of the robot is used to move or control the robot. Motor driver used is Smart Peripheral Controller DC Motor using the I2C-bus as a data delivery path that can save and simplify wiring. The robot recognizes the surrounding area make using sensors. There are various types namely ultrasonic sensors, infrared, GPS, cameras, accelerometers, and encoders. Distance sensor used is ultrasonic sensor provides a long distance measurement from 3cm- 3 meters. For navigation we use compass sensor to get start and goal position [9].

2.2. Tigal Voice Recognition System for Robotics

Voice recognition is the process of making the spoken word as an input for the computer program. Because there is a technology called the voice recognition, the authors created robot guidance in supermarkets are characterized as a human being, who has the ability to learn and know what humans said from sound. Most of the signals, such as biological signals, seismic signals, radar signals, and in particular is the sound is an analog signal. In order to process analog signals with digital tools cannot be directly processed. Surely, it must first be converted into a digital signal, which converts into a sequence of numbers that have limited precision. This procedure is called analog to digital conversion [5].

The Tigal SmartVR + DK-T2SI Voice Recognition Development Kit are for VoiceGP Module, which includes all of the hardware and software you need to develop voice and speech recognition capabilities into our robot. It includes the Sensory's Quick T2SI software to develop speaker independent vocabularies from text-based input in multiple languages with Sensory's revolutionary T2SIT technology. The VoiceGP module is a development platform for speech synthesis and voice recognition applications, based on Sensory RSC-4128 mixed signal processor.

Figure 2. TIGAL SmartVR Voice Recognition System [7]

This system implementing Hidden Markov Model (HMM) and having some terminologies:

1. Speaker Dependent: - systems that require a user to train the system according to his or her voice.
2. Speaker Independent: - systems that do not require a user to train the system i.e. they are developed to operate for any speaker.
3. Isolated word recognizers: - accept one word at a time. These recognition systems allow us to speak naturally continuous.
4. Connected word systems allow speaker to speak slowly and distinctly each word with a short pause *i.e.*, planned speech. Spontaneous recognition systems allow us to speak spontaneously.

3. Proposed Systems

Speech Recognition technology promises to change the way we interact with machines (robots, computers *etc.*,) in the future. We designed a prototype of robot guidance using distance sensor, compass, microcontroller, motor, driver motor and voice recognition system as shown in Figure 3 below. To create robot guidance in supermarket, we use low cost microcontroller, compass sensor, driver and DC motors and voice recognition system. Our system better than previous work such as [10], where we don't use IR for transmitting data.

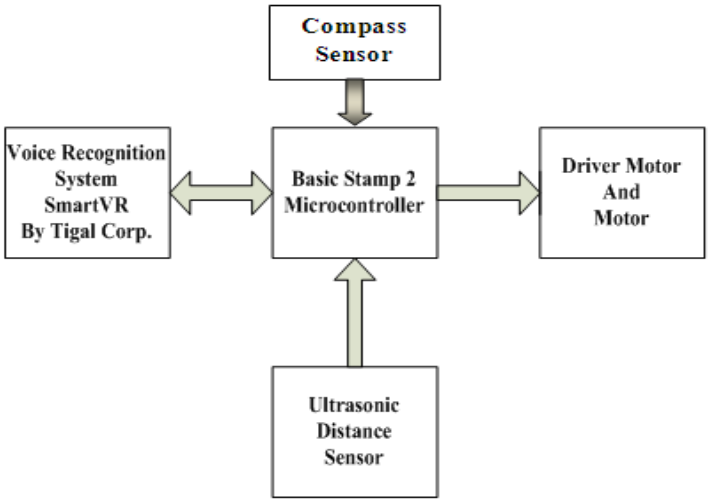


Figure 3. Architecture of Robot Guidance

The framework is shown in Figure 4, user will be prompted to enter a trigger word to activate your robot. After the user mentioning trigger word, robot will be active and welcoming user who comes to the supermarket. After robot giving welcoming words, the robot will ask what the user is looking for. Users can mention what objects are in looking to finish your robot after robot asked. When the user enters the command word the word robot will process and will take the user to the area of items sought.

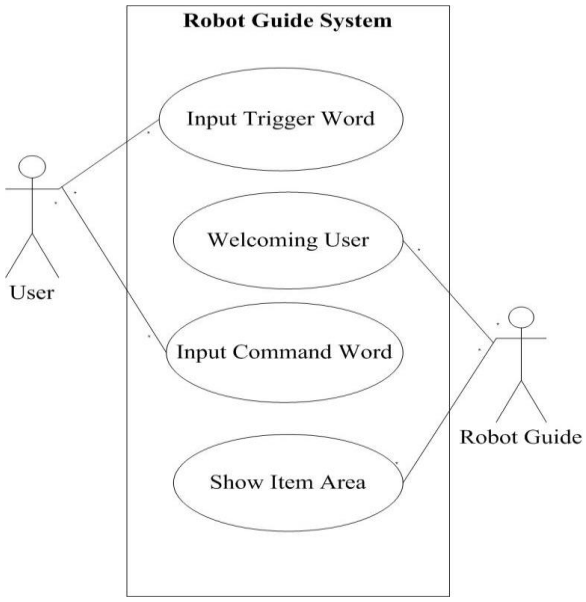


Figure 4. Use Case Diagram for Robot Guidance

Figure 5 shows a class diagram that describes a class diagram of robot guidance. There are several classes in the class diagram is the guest class, class robot, microcontroller class, class distance sensors, voice recognition class, and class of motor drivers.

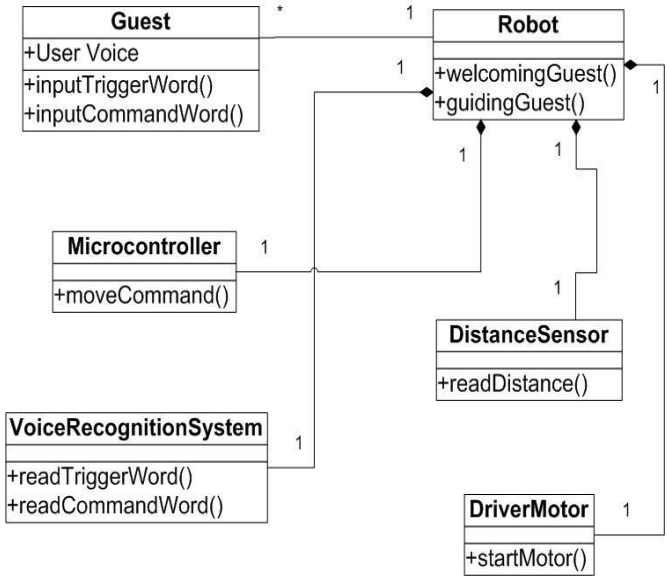


Figure 5. Class Diagram for Robot Guidance

Microcontroller is an integrated circuit that contains many items that are similar to those of a desktop computer, such as CPU, memory, I / O, *etc*.. Micro controller is a robotic brain. With the micro-controller allows the designer to adjust the sensor and contains the entire logic robot. Logic can be programmed using many languages. The microcontroller used is BASIC Stamp microcontroller [8]. Voice recognition is the process of making the words are pronounced as input for a computer program [7]. Voice Recognition System used is SmartVR by Sensory RSC-4128 mixed signal processor. The size is very small (42 x 72 mm) and has 2 connectors at the edges with 2.54mm pin spacing.

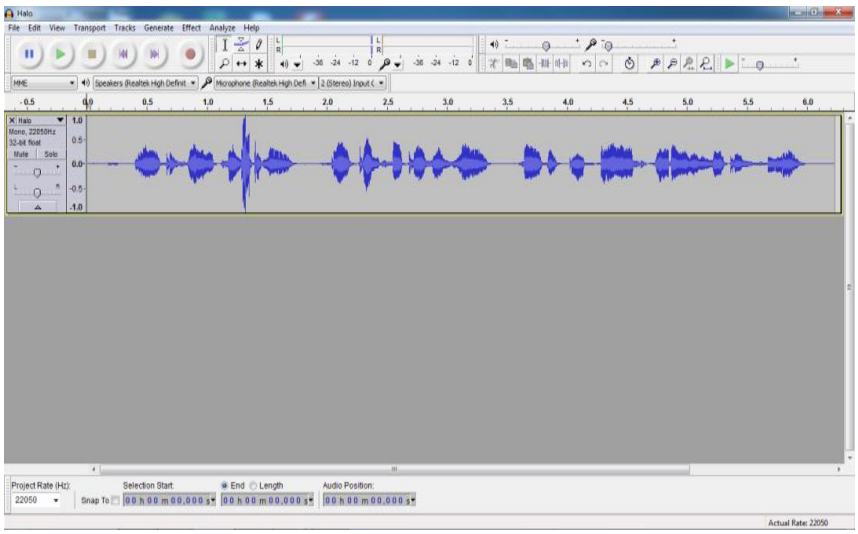


Figure 6. Example of Audio Signal for Welcoming the Customer from Google Voice and Processed using Software AudaCity

To adjust trigger word for the robot, we use T2SI (Text to Speaker Independent).

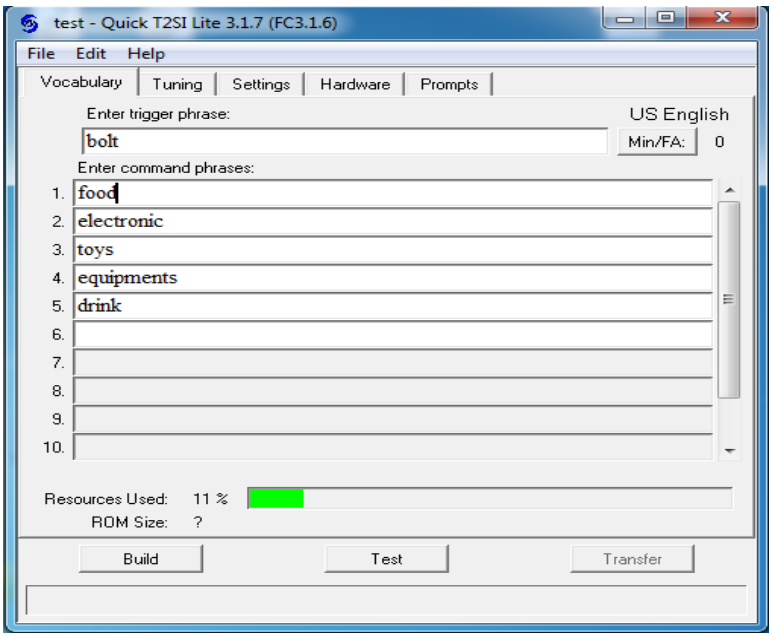


Figure 7. TIGAL SmartVR Voice Recognition System for Adding Commands [7]

For tuning the trigger word, we set the pronunciation as shown below:

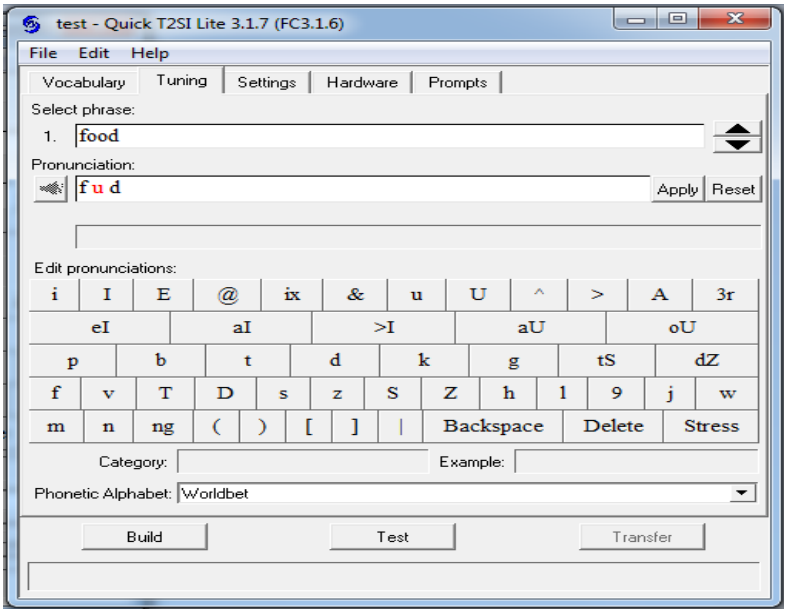


Figure 8. T2SI for Tuing the Trigger Words[7]

Figure 9 shows the flowchart of our proposed system, consist of input trigger word, welcoming customers, accepting input commands and read command word to navigate the robot to the goal position.

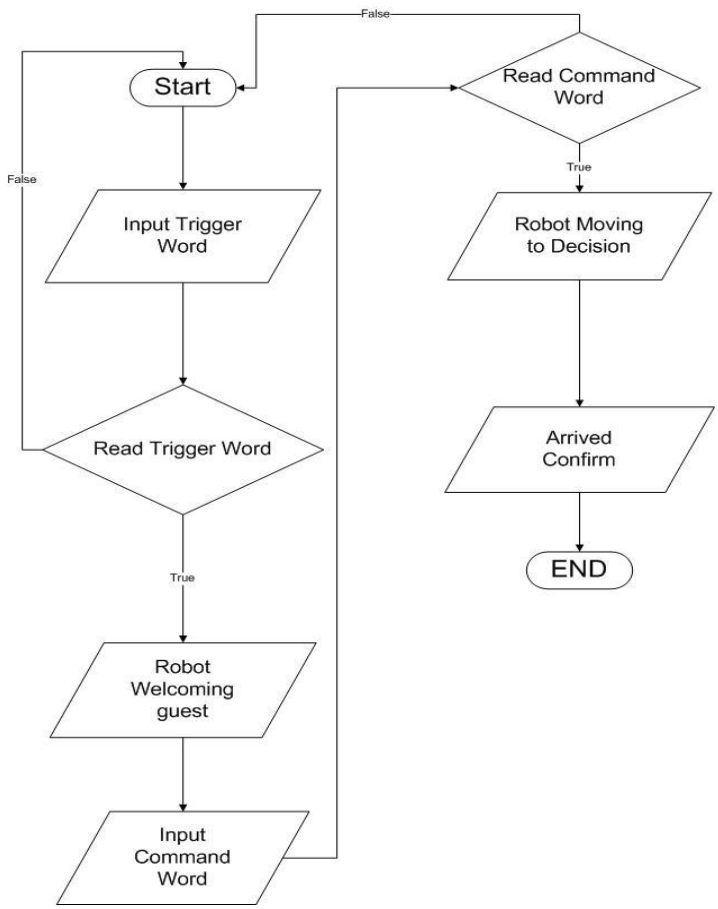


Figure 9. Robot's System Flowchart

Users can say a trigger word that serves as a caller to the robot. At that time, voice recognition system will attempt to identify the trigger word. If the robot does not recognize it, the robot will not respond to the user. If the robot recognize what was pronounced by the user as a trigger word robot will respond to the user with a sentence or sound output to welcome users who say the trigger word and the robot will ask you about the area of item location that user is looking. In this process the user will say a command word or area of items that the user is looking, if the robot does not it, then the robot will say sound output that the robot does not recognize the word and the user needs to repeat the pronunciation of the initial process. But if the robot recognized the command word spoken by the user, the robot will receive the command word and show the area location of item that wanted by the user.

4. Experimental Results

A test-plan is an important part of a testing session. It gives us an outline for testing and evaluating the system. We have designed a test-plan for our system testing and according to this test-plan have done our testing on the system. The system had been tested in the laboratory and outside in order to detect the environment effect on the recognition rate. Figure 8 shows our prototype of the robot:

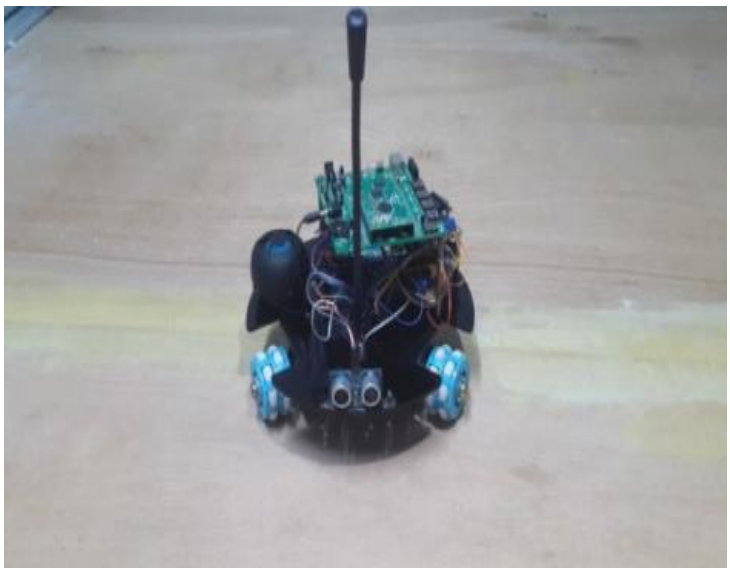
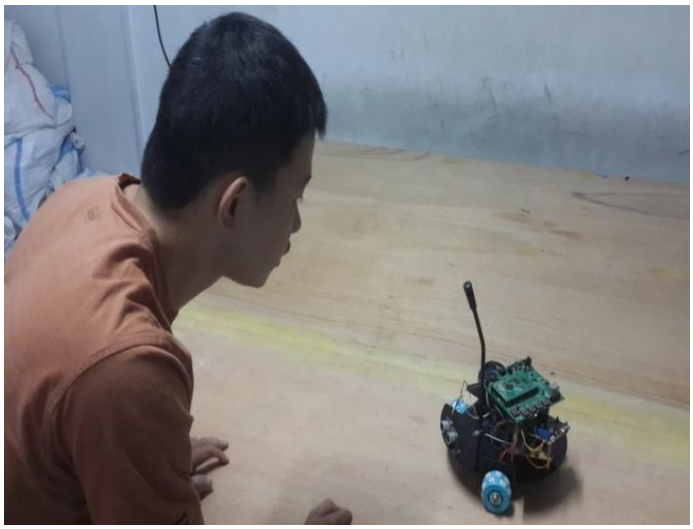


Figure 10. Robot Guidance for Store/Supermarket

The experiment conducted by giving some commands by varying distance of the microphone from the speaker and considering the rate of recognition with low noise and with higher noise as shown in Figure 11.



Firuge 11. Testing the Robot Guidance

Testing is conducted to test the overall system using voice from a speaker. After doing some testing with the English word, which is done by some speaker on voice recognition systems. From the test results we have found that the Module's speaker independent feature is enough sensitive. To get better speech recognition result -in our case to control the robot activities, you have to maintain the same distance with microphone and also the same tone. Table 1 shows the results:

Table 1. Test Result for Voice Recognition System

No	Word	Testing		Result
		T	*F*	
1	FOOD	7	3	70%
2	ELECTRONIC	9	1	90%
3	TOYS	8	2	80%
4	EQUIPMENT	8	2	80%
5	DRINK	8	2	80%

Table Description:
T = True F = False
Total words = 5 words
Total of repetition = 150 times
Recognized words = 120 words
Total accuracy = 80% accurate

5. Conclusion

Human-Robot interaction is an important, attractive and challenging area in HRI. The proposed system consists of a microcontroller and a voice recognition processor that can recognize many voice patterns. The conclusions to be drawn from this research are:

1. Robot guidance has a voice recognition system with pretty good accuracy with a total accuracy of 80%. The robot able to show the area to find the right item according to the accuracy of every word used in voice recognition with a total accuracy of 80%.
2. The robot can stop if there are obstacles in front of the sensor. For development and further research, the authors provide advice to its readers so that in the development of voice recognition system on the robot guidance in supermarkets could do more and better for the future.

Suggestions that can be provided are by adding or increase area list, more items, and make the area wider and get the location of items that wanted by customer clearer.

References

[1] A. Ronald, "Behavior –Based Robotics", The MIT press, Cambridge, Massachusetts, UK **(1998)**.
[2] M. Fezaria nd M. Bousbia-Salah, "A voice command system for autonomous robots guidance", 9th IEEE International Workshop on Advanced Motion Control, pp. 261 - 265, Istanbul, **(2006)**.
[3] J. Baumann, "Voice Recognition", URL http://www.hitl.washington.edu/scivw/EVE/I.D.2.d. VoiceRecognition.html.
[4] S. Y. Kavita and M. M. Mukhedkar, "Review on Speech Recognition", (IJSE) International Journal Of Science and Engineering, vol. 1, no. 3, **(2013)**, pp. 61-70.
[5] S. Preeti and K. Parneet, "Automatic Speech Recognition: A Review", International Journal of Engineering Trends and Technology, vol. 4, no. 2, **(2013)**.
[6] C. Schuler and C. Mahesh, "Digital Signal Processing: A Hands-On Approach", McGraw-Hill, **(2005)**.
[7] http://www.veear.eu/
[8] Anonymous, (2013). What's a Microcontroller?, http://parallax.com.
[9] T. Braunl, "Embedded Robotics: Mobile Robot Design and Applications with Embedded Systems", 3rd Edn., Springer Publisher **(2008)**.
[10] B. brahim, F. Mohamed and B. Mouldi, "A Voice Guiding System for Autonomous Robot", Asian Journal of Information Technology, vol. 4, no. 12, pp. 1201-1207, **(2005)**.

Author

Widodo Budiharto, Widodo Budiharto is a senior lecturer at School of Computer Science in Bina Nusantara University, Jakarta-Indonesia. He got PhD From Electrical engineering, Surabaya Institute of Technology at 2012. His research areas are Artificial Intelligence and Embedded systems.