International Journal of Research in Advent Technology, Vol.4, No.12, December 2016
E-ISSN: 2321-9637
Available online at www.ijrat.org

LOPEZ: A Bilingual Robotic Car

Kuo-pao Yang[1], Nicholas Moran[2], Isabella Bendana[3], Seth Champagne[4], Tyler Becker[5]
Department of Computer Science and Industrial Technology[1, 2, 3, 4, 5]
Southeastern Louisiana University[1, 2, 3, 4, 5]
Email: kyang@selu.edu[1]

Abstract- This paper describes a case study for a robotic car with speech recognition capabilities. LOPEZ is a bilingual robotic car, which can understand both English and Spanish. This speech recognition project uses CMU Sphinx, an open source project, to map audio recordings into control commands for the robotic car. This application runs as a server on a Raspberry Pi, which is mounted on the robotic car. The commands can be recorded on a computer and streamed to control the robotic car through wireless connection.

Index Terms- Robotic Car, Speech Recognition, Raspberry Pi.

1. INTRDUTION

Human-robot interaction with machines using such as hand gestures, eye motion, and voice controls has been proposed and studied for many years [10][11]. There are many technical and practical challenges for the development of the human-robot interaction components.

LOPEZ is a robotic car with speech recognition capabilities. This speech recognition system hosts a web site to receive dictation, understand, and carry out spoken commands. Voice commands can be recorded for the mobile robot execution. A voice control node module is used to receive audio and render it for commands to control the robot. This project currently supports multiple languages.

LOPEZ is a bilingual robot because it is able to recognize and interpret both English and Spanish commands. This system is able to not only understand instructions in English, but also respond to Spanish commands. For example, a user says "Go" to LOPEZ and the robotic car will move forward a few feet. The user can also say, "Vas" and LOPEZ will perform the same command.

This speech recognition project uses CMU Sphinx [3], an open source project from Carnegie Mellon University, to recognize audio recordings. Speech recognition is the ability of a machine or program to identify words and phrases in spoken language and convert them into a machine-readable format. CMU Sphinx [5] software is one of speech recognition applications. There is a difference between voice recognition (recognizing who is speaking) and speech recognition (recognizing what is being said). Using the knowledge base of speech recognition, the audio recordings are mapped into text format commands by CMU Sphinx and are executed on the robotic car.

The client-server architecture enables LOPEZ to provide a web service. This application is a server that runs on Raspberry Pi [2], which is mounted to the robot car. This system architecture is shown in Figure 1. A user on a client machine can control the robotic car through a web browser. If the user allows access to the microphone on the client computer, the robotic car can follow commands by clicking on control buttons or by recording voice settings.

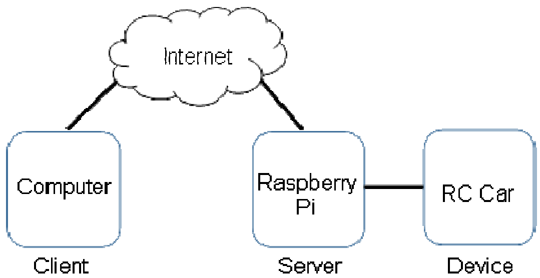

Fig. 1. LOPEZ System Architecture

Some work related to human robot interaction is described here. The VoiceBot is a voice controlled robot arm [7]. Human voice can specify control signals to manipulate the 3D robotic arm. The NavChair [12] is a voice controlled wheel chair that paraplegics use by dictating their direction. This voice control system can operate a power wheelchair with navigation assistance, which uses sensors to identify and avoid obstacles in the wheelchair's path. The PBDR [8] has been developed as a dance partner robot. After listening for the beats of music, the robot dances a waltz as a female dancer along with a human male dancer. Using cellular phones, the mobile robot [9] has the speech recognition capability to understand voice commands spoken in a natural way.

This paper presents a particular study for a robotic car with speech recognition capabilities. The related research in the current field for voice control to robotic cars is introduced. Both software and hardware approaches for full implementations are developed and detailed in this paper. Finally, conclusions are presented and issues for future research are discussed.

2. RELATED WORKS

Speech recognition for use with robotic control continues to be an advancing and promising field. Rudimentary speech recognition software has a limited vocabulary of words and phrases and may only identify these if they are spoken very clearly. More sophisticated software can accept natural speech. The process of speech recognition is that a computer receives audio inputs and then translates them into text for the computer to understand the voice commands. This text can be as simple as a small phrase, a very complicated statement, or a sentence that the computer can use to perform some tasks.

There are many voice controlled systems that exist. Users can have voice control through their cell phone. Apple Siri [1][4] is an intelligent virtual assistant that can understand voice commands and perform tasks like sending messages, placing calls, and making dinner reservations. But, Siri only understand one language at a time. If the user would like to use Spanish, the user needs to switch Siri language settings to Spanish.

Computers also implement speech recognition to be able to perform tasks. Windows Cortana [13][14] is a digital personal assistant used to perform many tasks through voice commands. Cortana can execute user voice commands to open Windows Explorer to browse files, send emails to users, and open Window Media to play video clips. What makes Cortana special is her ability to be able to hold conversations with the user. Cortana can respond with jokes, and other witty remarks to the user. But, she is not able to understand English and Spanish at the same time.

Google Translate [6] provides a multilingual translation service to translate text, speech, images, sites, or real time videos from one language into another. Text can be entered via a keyboard, handwriting recognition, or speech recognition. After entering searches in a source language, the user can select translation for the destination language. Google Translate supports over 100 languages at various levels and serves over 200 million people daily.

The advances in robotics technology have encouraged the development of areas in robotics and speech recognition. By using these new technology, LOPEZ presents a simple control system that uses speech recognition to wirelessly control a mobile robot. Voice commands used in this system are listened, interpreted, and then executed by the robotic car through the remote internet connection.

This speech recognition project currently supports multi-lingual service without switching language settings. There are many robots capable of responding to various language inputs. However, these types of human-robot interaction either provide only one native language or are limited to one language at a time. LOPEZ achieves multi-lingual options without having to switch languages.

3. IMPLEMENTATION

LOPEZ is a bilingual voice controlled robot, which is able to understand and execute spoken commands shown in Figure 2. These commands spoken in English or Spanish are registered, mapped, and converted to individual functions. Then, the robot follows these commands to drive left, right, forward, or backward. Without switching settings, LOPEZ searches the library of multiple languages to perform tasks. For example, a user says “right” or “east” in English to LOPEZ, the robotic car will turn right. The user can also say, “derecho” and LOPEZ will perform the same command.

```
var dictionary =  {
   deviceName  :  'lopez',
   commands :  {
      'right'     :  'right',
      'east'      :  'right',
      'derecha'   :  'right',
      'left'      :  'left',
      'west'      :  'left',
      'izquierda' :  'left',
      'forward'   :  'forward',
      'north'     :  'forward',
      'adelante'  :  'forward',
      'back'      :  'backward',
      'south'     :  'backward',
      'atras'     :  'backward',
   }
}
```

Fig. 2. Bilingual Voice Commands to Control LOPEZ Robot Car

LOPEZ includes two modes to operate the robotic car through wireless connection. The controlling mode allows the user to practice LOPEZ like a regular remote control car. It is a simple process that allows the user to test how to control and operate the robotic car. In Figure 3, LOPEZ provides a web page to accept the user inputs. The user can click on the Forward, Backward, Right, and Left buttons to control the robotic car.

In the recognizing mode, voice commands are interpreted and transferred into machine-readable format and then execute on the robotic car. A LOPEZ node module implements CMU Sphinx voice recognition software in a JavaScript environment. It allows a client computer to signal individual General Purpose Input Output (GPIO) pins on the Raspberry Pi. When the user clicks on the Record button, the computer will record the voice commands in an audio file called Lopez.wav. After the user clicks on the Stop button, the client computer sends the recorded wav file to the server for the robotic car execution.

The current method of controlling LOPEZ is through a client-server architecture. The Secure Shell (SSH) provides a secure channel for a remote terminal connection. Once connection to the Raspberry Pi is established, the user can run any applications under Raspbian operating system. LOPEZ is a speech recognition application, which hosts a web site to provide a user interface for basic robot controls and voice recording service. After the user enables LOPEZ to run as a web server through a SSH remote terminal connection, the user on the client computer can use any web browsers to control the robotic car through a wireless network.

Fig. 3. Webpage for LOPEZ

In recognizing mode, Raspberry Pi can receive voice commands. This system imports the library of the CMU Sphinx in English and Spanish language models to decode audio settings into text format. After decoding the audio, this system sends a corresponding text command to the server on the Raspberry Pi. The robotic car performs the desired commands by sending digital signals to a series of motor drivers.

For example, the user clicks on the Record button to record a voice command like "please go forward" in English. LOPEZ stores the audio command into a wav file. After the user clicks on the Stop button, the recorded wave file is sent the Raspberry Pi server for execution. LOPEZ calls the speed recognition library of the CMU Sphinx, which may return a text format like "go forward." This text format is mapped into the dictionary like "forward" in English. LOPEZ executes this command by calling the JavaScript functions to drive the GPIO pins on the Raspberry Pi.

The Raspberry Pi shown in Figure 4 comes with a 1.2GHz 64-bit quad-core ARM8 CPU, 1GB of RAM, 802.11n Wireless LAN, Ethernet Port, Micro SD card slot, HDMI port, and VideoCore IV 3D graphics core. The Raspberry Pi packages comes with a Micro SD card, which has a NOOBS operating system that allows to select either from the current list or get an operating system from the internet. This project chooses the Raspbian operating system, which is like Ubuntu. Raspabian is a free operating system based on Debian. Raspabian is optimized for the official Raspberry Pi hardware.

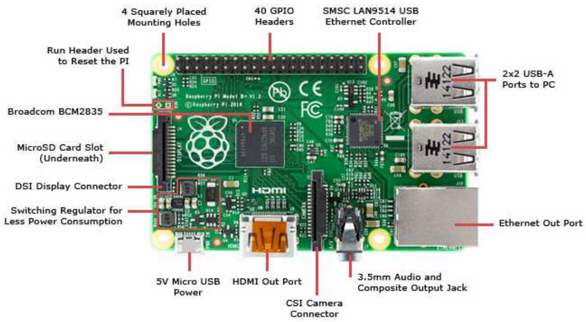

Fig. 4. Raspberry Pi

Figure 5 shows hardware connections for the robotic car. Other equipment mounted on the mobile robot includes two bread boards, four wheels, four motors, and a 4-AA power bank. These two bread boards are used for jumper cables and couplers to send signals to the respective motors. The whole robot is under three pounds when fully loaded. The chassis is made of aluminum, a very light metal to save the battery life. The motors are powered by a 4-AA battery pack hidden within the chassis. It creates a current through the bread boards which is then sent to each individual motor.

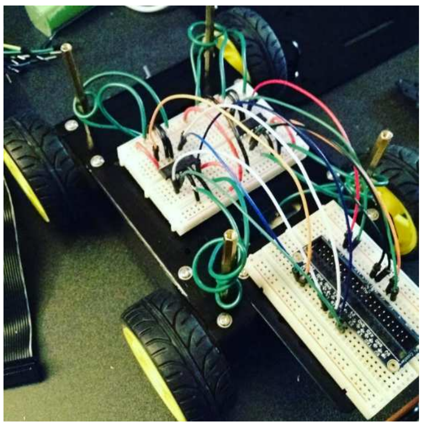

Fig. 5. Hardware Connections for the Robotic Car

The motors are controlled by the Raspberry Pi by sending signals to motor drivers through the GPIO pins. Figures 6 shows the GPIO pinout diagram of the Raspberry Pi. Each wheel has its own independent motor controlled by the Raspberry Pi. Each motor can move forward or backward at different speed like slow, medium, or fast. The robotic car turns by rotating wheels on one side faster than the other side. The robotic car drives straight forward and backward when all wheels are moving at the same speed.

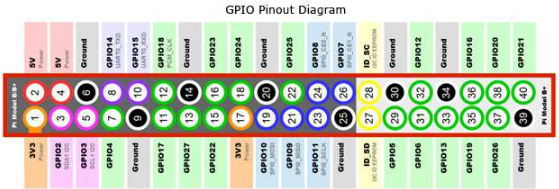

Fig. 6. Raspberry Pi GPIO Pinout Diagram

The finished product of the hardware implementation for LOPEZ is shown in Figure 7. The small black box on the top of the robotic car is the Raspberry Pi, which has a GPIO breakout broad with a black ribbon cable for connecting to a small bread board. This bread board then handles all the wiring connections. The Raspberry Pi is powered by a small power bank, which holds a capacity of 2400 milliamps. Its power is enough to keep the Raspberry Pi operating for at least 24 hours. This battery connects to the Raspberry Pi using a micro USB cable.


Fig. 7. LOPEZ: A Bilingual Robotic Car

This project employs CMU Sphinx software as a speech recognizer to transform speech input into written text. Testing the voice recognition involved in testing English and Spanish models. The results of the English voice model have 75% correct rate. The results of the Spanish voice model are correct in approximately 60% of the attempts.

4. CONCLUSION

LOPEZ is a speech recognition system, which hosts a web site to receive dictation, understand, and execute voice commands on a robotic car. The project currently supports both English and Spanish commands. LOPEZ includes two modes to operate the robotic car through wireless connection. The controlling mode allows use like a regular remote control car. In the recognizing mode, voice commands are interpreted and transferred into machine-readable format and execute on the robotic car. Improving the speech recognition will be the future work. We are working on Google Translate API to improve the speech recognition system.

REFERENCES

[1] Apple Siri, Apple Corporation, iOS - Siri - Apple, http://www.apple.com/ios/siri/, accessed December, 2016.

[2] Byrne, J., Fisher, L., and Tangney, B., "Computer Science Teacher Reactions Towards Raspberry Pi Continuing Professional Development (CPD) Workshops Using the Bridge21 Model," *10th International Conference on Computer Science & Education (ICCSE)*, 267 - 272, July, 2015.

[3] CMU Sphinx, Open Source Toolkit for Speech Recognition Project by Carnegie Mellon University, http://cmusphinx.sourceforge.net/, accessed November, 2016.

[4] Fourney, A., and Dumais, S., "Automatic Identification and Contextual Reformulation of Implicit System-Related Queries," *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval (SIGIR '16)*, 761-764, July, 2016.

[5] Ganesh, D., and Sahu, P., "A Study on Automatic Speech Recognition Toolkits," *International Conference on Microwave, Optical and Communication Engineering (ICMOCE)*, December, 2015.

[6] Google Translate, http://translate.google.com, accessed December 2016.

[7] House, B., Malkin, J., and Bilmes, J., "The VoiceBot: A Voice Controlled Robot Arm," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, 183-192, April, 2009.

[8] Kosuge, K., "Dance Partner Robot An Engineering Approach to Human-robot Interaction," *Proceedings of the 5th ACM / IEEE International Conference on Human-Robot Interaction (HRI '10)*, 201-201, March, 2010.

[9] Kubik, T., and Sugisaka, M., "Use of a Cellular Phone in Mobile Robot Voice Control," *Proceedings of the 40th SICE Annual Conference*, 106 – 111, July, 2001.

[10] Modares, H., Ranatunga, I., Lewis, F., and Popa, D., "Optimized Assistive Human-Robot Interaction Using Reinforcement Learning," *IEEE Transactions on Cybernetics*, 46(3): 655-667, March 2016.

[11] Rafsan, S., Arefin, S., Hasan, A., and Hoque, M., "Design a Human-Robot Interaction Framework to Detect Household Objects," *5th International Conference on Informatics, Electronics and Vision (ICIEV)*, 973 – 978, May, 2016.

[12] Simpson, R., and Levine, S., " Voice Control of a Powered Wheelchair," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 10(2): 122-125, November, 2002.

[13] Windows Cortana, Microsoft Corporation, Search for anything, anywhere - Windows Help. http://windows.microsoft.com/en-us/windows-10/getstarted-search-for-anything-cortana, accessed December, 2016.

[14] Ye, G., Liu, C., and Gong, Y., "Geo-location Dependent Deep Neural Network Acoustic Model for Speech Recognition," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6075 – 6079, May, 2016.