System Requirements Specification
CS308 project
Speech Based Servant Bot

Ganji Dhiraj Govind Chintapalli Gaurav Torka Sai Teja Pratap

March 1, 2012

Contents

Chapter 1

Introduction

We plan to create a **interactive remote servant robot** which takes the instructions from user in the form of speech via an android device (with the user). Thus one can control multiple bots by using a single device which passes on the instructions to bots.

The main purpose of this document is to give a detailed description of the Speech Based Servant Bot,explain the features , interfaces of the system and what the sysem does. It also explains the constraints under which it must operate and how the system will react to external stimuli.

1.1 Definitions,Acronyms and Abbreviations

- **Bot/Robot**The electro-mechanical machine that is guided by the user.
- **User** The person that is currently controlling the bot.
- **Sphinx** A java speech processing library
- **PocketSphinx** A version of Sphinx that can be used in embedded systems
- **Zigbee** A wireless communication protocol
- **Arm** The mechanical arm attached to the robot which lets the bot to pick up object.
- **Esterel** synchronous programming language for the development of complex reactive systems.

1.2 References

- http://cmusphinx.sourceforge.net/sphinx4/
- http://www.tbrk.org/esterel/index.html
- http://processors.wiki.ti.com/index.php/Android_ZigBee_Interface
- http://cmusphinx.sourceforge.net/2011/05/building-pocketsphinx-on-android/
- http://swathiep.blogspot.in/2011/02/offline-speech-recognition-with.html

Chapter 2

Overall Description

The bot acts as a servant which performs basic actions as specified by the user in the form of speech. The Android device processes the speech and generates instructions for the bot to execute.The user can ask the bot to go to a place from its position and send images of the new place based on which the user further instructs the bot.

2.1 System Environment

The environment has two active components and one coordinating system. The user communicates to the coordinator the instructions to be executed in the form of speech. The coordinator processes the input (speech) and sends instructions to the bot to perform via a dedicated wireless frequency channel.

2.2 Product Perspective

- The user specifies commands to the Android device in the form of speech.
- The Android device processes the input speech and interprets the commands to be executed by the bot.
- The speech processing is done with the help of the speech processing library(PocketSpinx)
- The processed commands are then remotely forwarded to the bot using the ZigBee wireless interface.
- The commands to the bot are translated to signals using Esterel.
- The bot receives the signals and acts accordingly.
- A camera will be mounted on the robot to allow the user to view the environment of the bot and the bot to see it its own environment.

2.3 Product Functions

- The bot can move around in the arena as specified by the user. It can move forward, backward, turn left, turn right, speed up, slow down, etc.
- The bot can be instructed to go to a remote place, send images of that place to the android device. The images can then be analysed by the user , based on which he/she can decide the instructions to be given to the bot.
- The bot can detect objects through image processing and perfom actions like picking up objects or dropping them as per instructions.
- A torch is attached to the bot, which can be used in dark places to capture the view.

2.4 User characteristics

The user should be familiar with the commands that can be executed. He/she is expected to be able to speak fairly consistently and with a good level of resemblance so that it can be detected properly by the Speech Library. The main screen of the Coordinator will have images of the environment captured by the bot.

2.5 Constraints

- For the speech library to function efficiently, we need to provide a decent level of training.
- The machine used as the coordinator requires a good processor.
- The distance between the user (with the Android device) and the bot must be within a certain limit restricted by the ZigBee wireless protocol.
- To make the image processing and communication between the bot and android fast the image size has to be small. To achieve this the number of colors in the environment has to be reduced.

2.6 Assumptions and Dependencies

- The bandwidth of ZigBee Communication should be able to handle the load of communications i.e. downlink - the commands from the Android device to the bot; and uplink - the images from the camera mounted on the bot to the Android device + Errors if any.
- There must be no interference in the form of noise due to other users or any other external factors since it may affect the efficiency of the speech recogniser.

2.7 Requirement Subsets

- Speech Recognition
- Transmission of Commands
- Interpretation of Commands
- Signal generation based on Commands
- Action performed based the Signals
- Capturing the image of the environment
- Response based on the environment of the Robot

Chapter 3

Details

3.1 Functionality

This bot acts as servant which takes commands from the user in speech format.The bot can detect objects which the user can ask it to pick up and drop in a specified location.It can move to remote locations as per the user's instruction and send the image of the remote location. From this the user can further instruct the bot as per the environment in the remote location.Thus the user can perfom tasks without actually being present in the location. If he has multiple bots at different locations then he can make the bots work at different locations parallelly using a single android device

3.2 Interfaces

3.2.1 User Interface

This is the part of the system that will be used to interact with the user which includes processing and recognising speech given by the user as well as displaying image of remote location sent by the bot.

3.2.2 Hardware Interfaces

The following are the physical equipment required to interact with the bot and make it work.

Android device

This receives the speech from the user through its microphone, processes it using PocketSpinx speech library. It sends the instructions to the bot using ZigBee protocol. It also receives the images sent by the bot and processes them. For the ZigBee to work, an additional xxxxxxxxxxx has to installed.

Firebird V

It receives and executes the instructions sent by the android device.

- A **camera** to be mounted on the bot, which captures the images. It is connected to the bot via a serial port.

- A **torch** to be mounted on the bot which lets the bot work in the dark.
- An **robotic arm** which lets the bot to pick and drop objects.
- A **mount rotor** which lets the arm, light ,camera to rotate 360 degrees

3.2.3 Software Interfaces

The following are the software requirements of the project:

PocketSphinx

This is a speech processing library. This has to be installed in the android device. For the library to function properly it has to be trained with the speech inputs.

Andriod OS

llllllllllllll—————————————————————————————————

ZigBee module for android

The ZigBee module provides a way to access the ZigBee hardware through its library functions. Thus, the android can communicate with the bot using this module.

3.2.4 Communication Interfaces

The project uses ZigBee wireless protocol for communication between bot and android device.

Chapter 4

Quality Control

4.1 Recognising speech input

Each speech input corresponds to a distinct action to be performed. So the speech processing has to be done properly

4.2 Identifying the object to be picked up

If the bot is instructed to pick up an object then the android has to process the image send by the bot to identify the object. This requires image processing to be done properly.

Chapter 5

Risk management

5.1 Speech ambiguity

It is probable that the speech is not properly understood by the system. In that case it might behave unexpectedly. *Fall back plan*: In this case it will give a warning instead of unexpected behaviour. This can be achieved by setting a proper threshold for accepting an input.

5.2 Image transfer delay

If the environment is very complicated, then the image can't be compressed properly, which results in large size of images. So there would be a lot of delay to transfer and process the images. *Fall back plan*: To avoid this the environment is kept very simple.