System Requirements Specification
CS308 project
Speech Based Servant Bot

Ganji Dhiraj Govind Chintapalli Gaurav Torka Sai Teja Pratap

March 3, 2012

Contents

Chapter 1

Introduction

We plan to create an **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

Bluetooth 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.

Android Android is a Linux-based operating system for mobile devices such as smartphones and tablet computers.

App An application performing specific functions.

Config File The file that maps speech instructoin and the signal to be sent.

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 do any of pre-defined actions which are configurable according to the application.

2.1 System Environment

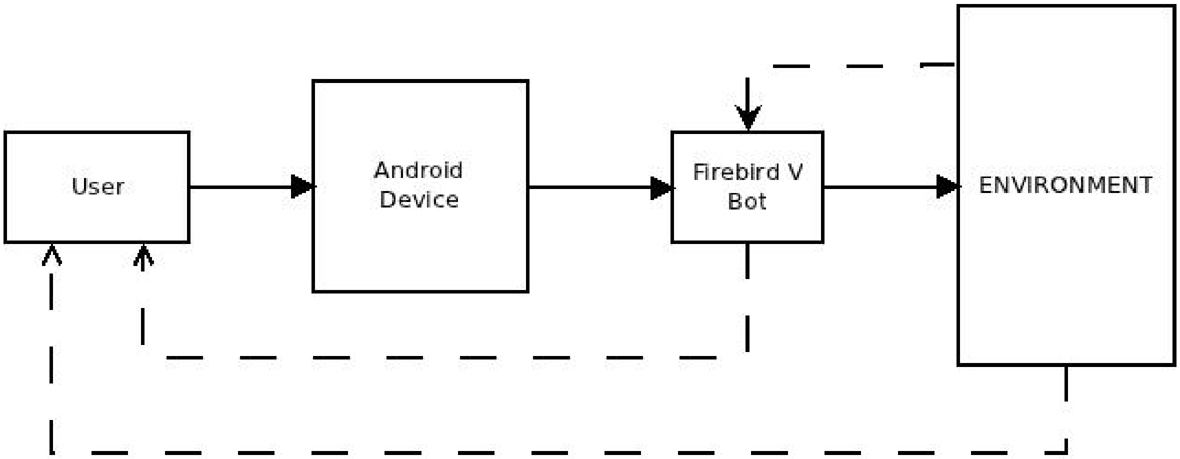

Figure 2.1: System environment

The system consists of the user, Android Device, Android App and the FireBird bot.

- The user speaks into the *Android device.*
- The *App* on the *Android Device* processes the input checks in the *Config File* and generates an action.
- The *App* sends the action to the *bot* over *Bluetooth.*
- On receiving the action, the *bot* executes the action according the *application.*

The environment if largely dependant on the application built upon this system. It could include white-lines, ping-pong balls and pretty red ribbons.

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 commands are translated to signals, and then remotely forwarded to the bot using the Bluetooth wireless interface.
- The bot receives the signals and acts accordingly.

2.3 Product Functions

The product should have to be interfaced with the *App*; i.e. it should understand the signals it receives from the *Android Device.* Similarly, the *App* should also be configured using the *Config File.* The *bot's* capabilities are dictated by the *application.* **e.g.**

- The bot can move around in the arena specified by the user.
- The bot has a physical arm which can pick and drop objects.
- The bot can detect objects through image processing.

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 coherently so that it can be detected properly by the Speech Library.

2.5 Constraints

- For the speech library to function efficiently, it needs to be trained.
- The distance between the user (with the Android device) and the bot must be within a certain limit restricted by the Bluetooth wireless protocol(10m).
- 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

- 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
- 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. 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.

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 PocketSphinx speech library. It sends the instructions to the bot using Bluetooth protocol.

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.
- 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

Android is a Linux-based operating system for mobile devices such as smartphones and tablet computers. It is developed by the Open Handset Alliance led by Google.

3.2.4 Communication Interfaces

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

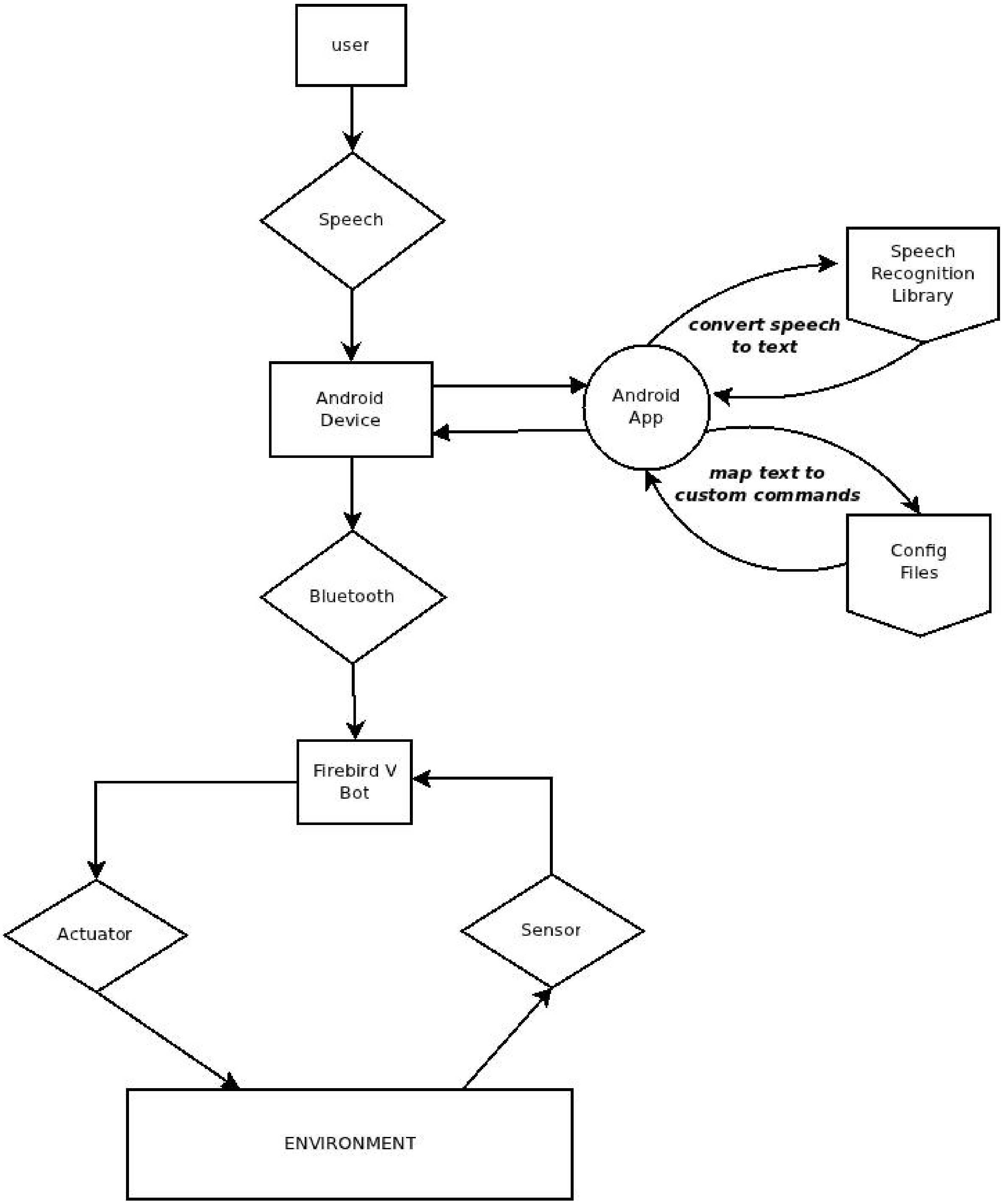

Figure 3.1: State diagram

Chapter 4

Quality Control

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

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 process 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.