

Voice based Application Designed for Visually Impaired People and enabling them to use Gmail Services

Rashmi Gupta

Dept. of Electronics and Communication Engineering
Ambedkar Institute of Advanced Communication
Technologies and Research
New Delhi, INDIA

Kaushal Sharma

Dept. of Electronics and Communication Engineering
Ambedkar Institute of Advanced Communication
Technologies and Research
New Delhi, INDIA

Deepak Chaudhary

Dept. of Electronics and Communication Engineering
Ambedkar Institute of Advanced Communication
Technologies and Research
New Delhi, INDIA

Rajiv Kapoor

Dept. of Electronics and Communication Engineering
Ambedkar Institute of Advanced Communication
Technologies and Research
New Delhi, INDIA

Abstract—Internet is flooded with the billions of users, in which they perform wide variety of tasks. But in this crowd, sometimes, visually impaired people feels left out in using some of basic and common mode of modern communication i.e. Gmail. So there must have to be a mechanism through which they can communicate like we do, through Gmail and in this paper, we are going to present the same using Python, Natural Language Processing, and Automation using Selenium.

Keywords— Python, Selenium, Natural Language Processing, Google voice recognition API and Google TTS API.

I. INTRODUCTION

The application which is presented in this paper is fully integrated with (runs with) voice commands only. The visually impaired people needs to give the voice command as input and the rest of the work is done automatically through automation code (Selenium).

There are wide varieties of other API's (Application Program Interface), and technologies which is used for taking the voice input through user and to convert it to the understandable form is Google Voice Recognition API, and for output purpose the text to speech converter is used i.e. Google TTS API, both these works on the principle of Natural Language Processing, when all the commands are clear to the application, then the rest of the work is being performed by the automation software, Selenium.

II. RELATED WORK

Some major work is already done for the voice commands and other stuffs for visually impaired people. Some of those are highlighted below:

- 1) Design and Implementation of Text & Speech conversion for visually impaired people (International Journal of Applied Information System - ISSN: 2249-0868) Author: Itunuoluwa Isewon, Jelili Oyelade and Olufunke Oladipupo [1].

- 2) www.ijarccce.com/upload/2015/january/IJARCCCE5C.pdf

III. OVERVIEW OF DIFFERENT TECHNOLOGIES USED Before

A. Google voice recognition API [3]

It converts the voice to text using neural networks, natural language processing, etc. Basic flow is given below:

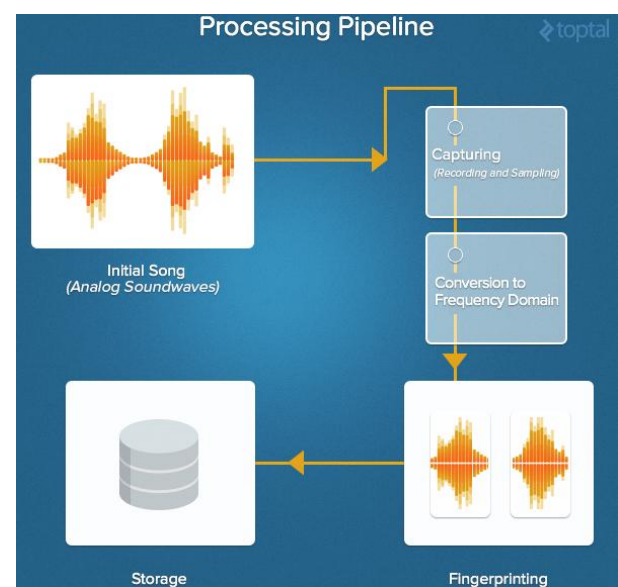


Fig. 1. Flow

B. Google Text to Speech API [4]

It converts text to speech. It also uses the natural language processing to tokenize the text and give considerable amount of pause (while speaking) after completion of a word.

C. Selenium [5]

It is a Software-testing/automation framework for the web applications and is an open-source software. We can use Selenium webdriver to surf the internet/automate things using python, java, etc. without the interference of human.

IV. BLOCK DIAGRAM

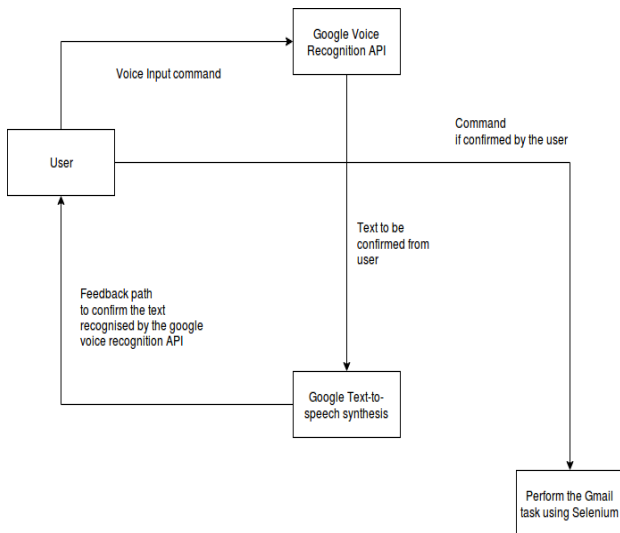


Fig. 2. Block diagram

V. FLOW CHART

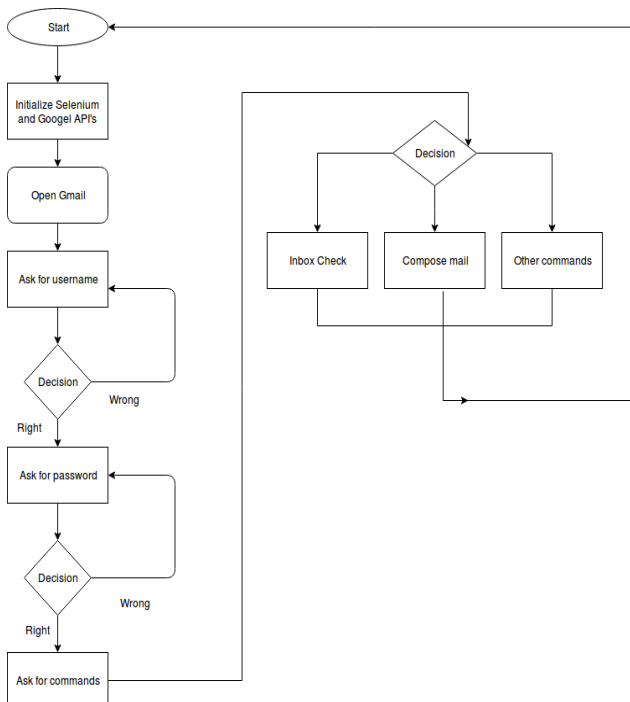


Fig. 3. Flow chart

VI. CODE SNIPPETS AND IMPLEMENTATION

The application is being developed in python and is hosted on Github (github.com/deepakchaudhary7/Voice-based-Gmail-system-for-visually-impaired-people/tree/master)

The various code snippets of the setup part are as follows

A. Setting up the selenium

```
from Selenium import webdriver
browser = webdriver.Chrome(chromedriver)
browser.get("https://www.google.com/gmail")
```

B. Setting up the google voice recognition API

```
import speech_recognition as r
With microphone() as source:
    r.adjust_for_ambient_noise(source)
    audio = r.listen(source)
return r.recognise_google(audio)
```

C. Setting up the Google Text to Speech Synthesis

```
from gtts import gTTs
tts = gTTs(text= text, lang = hi);
tts = save('voice.mp3)
os.system("mpg321 voice.mp3")
```

D. Snippet to execute the compose mail

```
browser.find_elements_by_xpath("//*[@id=':i8']/div/div)[0].click()
WebDriverWait(browser,15).until(EC.element_to_be_clickable((By.XPATH '//*[@id=':125'])))
speak('who m do you want to compose mail')
Receiver = voice_input()
speak('what is the subject')
Subject = voice_input()
speak('what is the message')
Message = voice_input()
browser.find_elements_by_xpath("//*[@id=':12w']")[0].send_keys(Receiver)
browser.find_elements_by_xpath("//*[@id=':12f']")[0].send_keys(Subject)
browser.find_elements_by_xpath("//*[@id=':13g']")[0].send_keys(message)
browser.find_elements_by_xpath("//*[@id=':125']")[0].click()
```

VII. FUTURE IMPROVEMET SCOPE

Although this is a good application for visually impaired people for using Gmail, but things can get challenging at times when noise is at the peak in the surroundings. So in future may be we can develop a system to rectify the noise while taking the input from the user. Security is also a big challenge and this can be easily improved using some pre-authentication techniques so that user doesn't have to speak password again and again.

REFERENCES

- [1] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.635.4588&rep=rep1&type=pdf>
- [2] <https://www.ijarccce.com/upload2015/january/IJARCCCE5C.pdf>
- [3] http://www.mibqyyo.com/en-articles/2016/04/14/how-does-google-voice-recognition-work/#/vanilla/discussion/embed/?vanilla_discussion_id=0
- [4] https://github.com/Uberi/speech_recognition

[5] [https://en.wikipedia.org/wiki/Selenium_\(software\)](https://en.wikipedia.org/wiki/Selenium_(software)).