Virtual Assistant: A Review

Kartik Chauhan[*]

Researcher & Software Developer, Optiwise IT, Dehradun, India
*Corresponding author: kartikchauhan100@gmail.com

Abstract: As we know that, virtual assistant can do anything as an intelligent system with the help of natural language processing. Our research paper is focus on to know virtual Environment and virtual Assistant Interfaces works, and the paper presents applications of virtual assistant that helps in accessing the software without having knowledge of operation of software especially in bilingual. It also describes limitation and the challenges that occur in virtual Assistant technology.

Keywords: Automation, Face recognition, Machine Learning, Motion detection, NLP, Python socket, Speech recognition.

1. Introduction

Virtual assistant is software that can do most of the things for the help of user and can save time. It can respond to all sorts of voice commands, send text messages, make phone calls, set up reminders; anything you do on your computer, you can probably ask your virtual assistant to do for you.

Virtual Assistant is a dynamic technology. It means that the computer is able to detect user voice command and give the best suitable and satisfied result to the user. Virtual Assistant with a focus on user-based information. It will look at examples of intelligent programs with natural language processing that are currently available, with different categories of support, and examine the potential usefulness of one specific piece of software as a Virtual Assistant. This engages the ability to communicate socially through Natural Language Processing, holding (and analyzing) information within the context of the user.

2. Work Done

Our project runs in four phases as follows:
Phase 1: Face Detection &Recognition
Phase 2: Speech to text
Phase 3: Text analysis
Phase 4: Interpret command
Phase 5: Text to speech

A. Phase-1 (Face Detection & Recognition)

The first phase of project is to unlock the virtual assistant using face lock. We are using face detection & recognition technique for creating face lock feature. We are detecting faces from image using face detection algorithm and we are matching those faces with each other with the help of face recognition algorithm. If face recognition algorithm predicts that 90% of face is similar, then it will unlock the virtual assistant.

Confidence value 0 denotes 100% face matching. We are using Standard confidence value of 37 for unlocking software. Means if the confidence value of face is less the 37 that means face is matched.

B. Phase-2 (Speech to text)

Second phase of our project is speech to text conversion in which the user gives voice command to the software using local language (English or Hindi). Then the system converts it into the text using speech recognition library then the extracted text is sent for text analysis.

C. Phase-3 (Text Analysis)

After the speech verification the software convert the user command into the computer understandable code, it is easy to access the data from the software.

D. Phase-4 (Interpret Command)

After converting the command into the computer understandable language, the virtual assistant check whether the command exist or not if command exist then software will perform operation otherwise the software will accept command as a question and search answer in database if answer found then the answer is given by software otherwise it save that question for future reference and recognize the next user command.

E. Phase-5 (Text to speech)

The last step of the software is to convert the text that is collected from database or internet into the speech to give the predicted answer to the user in local language.

3. Functions

A. Voice Detection

Steps:

1) Speaker Identification: Speaker Identification is a process to understanding who is speaking.

2) Speaker Verification: Speaker Verification is the process of accepting or rejecting the identity claim of a speaker Speech Recognition

3) Speaker Recognition: Speaker Recognition is a process to identifying what is said by the end user.

International Journal of Research in Engineering, Science and Management
Volume-3, Issue-7, July-2020
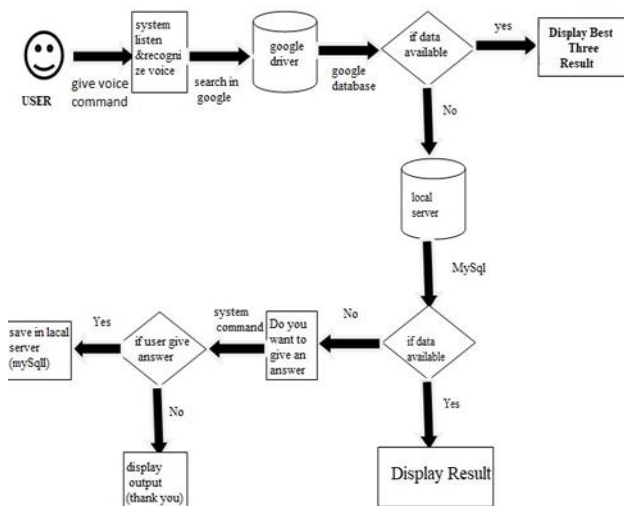**journals.resaim.com/ijresm | ISSN (Online): 2581-5792**

139
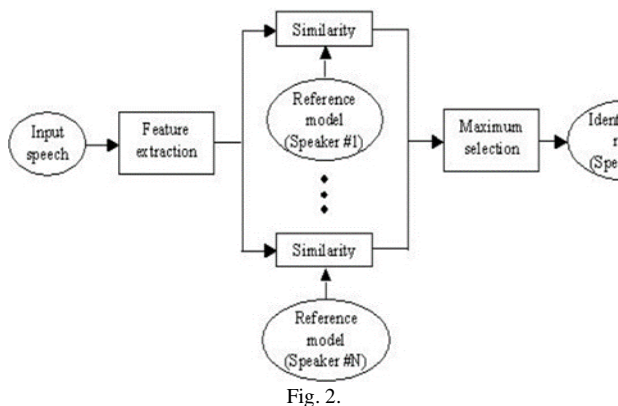
Fig. 1. Working process

Fig. 2.

B. Motion detection

This python program will allow you to detect motion and also store the time interval of the motion.

Before starting the actual motion detection capture the initial frame and match all the other frames with that initial frame. If there is difference in other frame as compared to initial frame, then motion will be detected and virtual assistant will inform user that someone entered in his room.

This feature only activates when user ask virtual assistant to monitor his/her room.

Requirements:
1. Python3
2. OpenCV(libraries)
3. Pandas(libraries)

1) Install Requirements

Install Python3, install Pandas and OpenCV libraries. Videos can be treated as stack of pictures called frames. Here I am comparing different frames (pictures) to the first frame which should be static (No movements initially). We compare two images by comparing the intensity value of each pixels.

C. Automation

Python is extremely user-friendly for interacting with just about anything.

Once you've mastered the basics of programming, you can create Python programs that effortlessly perform useful and impressive feats of automation to:
- Search for text in a file or across multiple files
- Create, update, move, and rename files and folders
- Split, merge, watermark, and encrypt PDFs
- Send reminder emails and text notifications
- Fill out online forms
- Update and format data in Excel spreadsheets of any size
- Search the Web and download online content

Steps:
1. Import the libraries.
2. specify the video to play.
3. Look for the YouTube search results.
4. open the video in the browser.

D. Sockets

Sockets and the socket API are used to send messages across a network. They provide a form of inter-process communication (IPC).

Python sockets is the python library which is used to send data on a network. Use python sockets to create a server on the network which accept data from the clients. And send data to clients. And by using Android Sockets Create two clients.

One client is used to monitor activities in Android smartphone and report it to the server and virtual assistant will inform us about any activity like call or message incoming in our smartphone.

Second client is used to send task to the server which is then completed by our virtual assistant. This whole process is a real time process.

4. Limitations

1. Virtual assistant can't speak Hindi language properly.
2. At some points main thread get paused because of some operations. Example: while creating web view.
3. Virtual assistant can only perform limited operations.
4. While automating Google chrome sometimes selenium driver gets disconnected because of ads on chrome.
5. Internet connection of minimum 2 Mbps speed is required otherwise performance of virtual assistant should decrease.

5. Conclusion

Our system can respond to all sorts of voice commands, send text messages, make phone calls, set up reminders; anything you do on your phone, you can probably ask your virtual assistant to do for you.

The main services provide by our project are as follows:
1. Provide system admin facilities like system shut down system sleep and Hibernate.
2. Music play form jio savan website.
3. Search jokes (Hindi or English).
4. Alarms.
5. Set remainder

6. Face lock.
7. Google search.
8. Voice detection.
9. Automation
10. Motion Detection

The benefit of our project is that anytime we use while we have an active internet connection, and it often self-employed or work for another body, often for multiple clients simultaneously, removing any need for you to provide benefits, organize payroll or taxes.

6. Future Scope

- We can use virtual assistant in home automation.
- We can connect virtual assistant with android so that we can automate computer and all devices connected with virtual assistant with android app.

- We can automate whole Google chrome using this virtual assistant.

References

[1] "Accurate and compact large vocabulary speech recognition on mobile devices," in INTERSPEECH. 2013, pp. 662–665, ISCA.
[2] "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," in ICML, 2006, pp. 369–376.
[3] "Lattice-based optimization of sequence classification criteria for neural-network acoustic modeling".
[4] "Language modeling for What-with Where on GOOG411," Proc. of INTERSPEECH, 2009.
[5] "Direct construction of compact context-dependency transducers from data," Proc. of INTERSPEECH, pp. 218–221, 2010.
[6] "Automatic question generation for decision tree-based state tying," Proc. of ICASSP, 1998.
[7] "Speech recognition with flat direct models," IEEE Journal of Selected Topics in Signal Processing, 2010.