

Survey on Virtual Assistants

Aniket Rode¹, Aditya Dudhakawar², Aniketa Maharana³, Prof. V. R. Surjuse⁴

^{1,2,3}Student, Dept. of Computer Technology, KDK College of Engineering, Nagpur, India

⁴Professor, Dept. of Computer Technology, KDK College of Engineering, Nagpur, India

Abstract - In today's world, Artificial Intelligence (AI) has become an integral part of human life. There are many applications of AI such as Chatbot, network security, complex problem solving, assistants and many such. Artificial Intelligence designed to have cognitive intelligence which learns from its past experience to take future decisions. A virtual assistant is also an example of cognitive intelligence. Virtual assistant implies the AI operated program which can assist you to reply to your query or virtually do something for you. Currently, the virtual assistant is used for personal and professional use. Most of the virtual assistant is device-dependent and is bind to a single user or device. It recognizes only one user. Our project proposes an Assistant that is not device bind. It can recognize the user using facial recognition. It can be operated from any platform. It should recognize and interact with the user. Moreover, virtual assistants can be used in a different area of applications such as education assistance, medical assistance, vehicles and robotics, home automation, and security access control.

Key words: Artificial Intelligence, Cognitive Intelligence, Virtual Assistant, Facial Recognition, Chatbot

1. INTRODUCTION

AI defines as any device that understands its surroundings and takes actions that increase its chance to successfully accomplish its goals. AI as "a system's ability to precisely interpret external data, to learn from such data, and to use those learnings to accomplish distinct goals and tasks through supple adaptation." Artificial Intelligence is the developing branch of computer science. Having much more power and ability to develop a various application. AI implies the use of the different algorithm to solve various problems. The major application being Optical character recognition, Handwriting Recognition, Speech Recognition, Video Manipulation, Robotics, Medical Implementation, Virtual Assistant, etc.

Considering all the application, Virtual assistant is one of the most influencing applications of AI and lately attracting the interest and curiosity of researchers. The virtual assistant supports a wide range of application and because of this it is categorized into many types such as virtual personal assistant, smart assistants, digital assistant, mobile assistant or voice assistant. Some of the well-known virtual assistant being ALEXA by Amazon, Cortana by Microsoft, Google Assistant by Google, Siri by Apple, Messenger 'M' by Facebook. These companies used different ways to design and improve their assistants. There are many techniques used to design the assistants based on the application and its complexity.

For example, Google uses the Deep Neural Network (DNN) for its components. Also, Microsoft uses the Microsoft Azure Machine Learning Studio to develop Cortana's components. However, their potential is limited by some scathing security issues that they don't support powerful authentication mechanisms and they are bind to their specific hardware. Face recognition or other identification mechanisms should be required before accepting any voice commands and they should not bind to any specific hardware.

In this paper, we propose an approach which will overcome the security issue with the help of Face and Speech recognition and using browser-based assistant it will overcome the hardware dedicated problem.

2. BLOCK DIAGRAM

Virtual Assistants is one of the active areas that many companies are trying their hands on to improve its efficiency and applications. There are many techniques are used to design the virtual assistants based on its application and complexity and there are many different architectures for it. Based on this data we designed a data flow diagram for assistant.

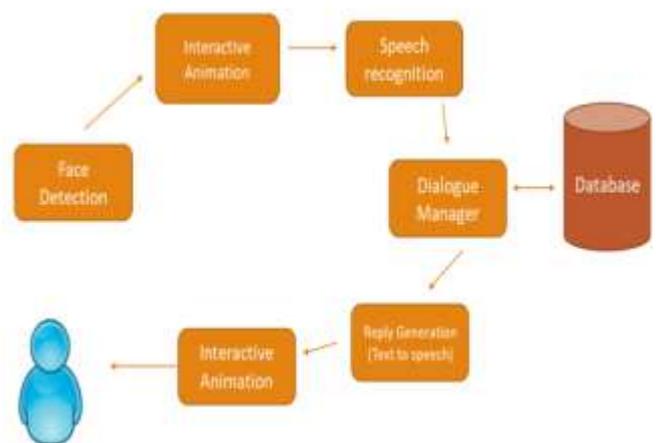


Fig-1: Data-Flow Diagram

Data flow diagram of the Interactive Animated Virtual Assistant is shown in Fig.1.1. It shows the flow of data from the user to the AI and the generation of the reply.

3. WORKING MECHANISM

This assistant is fully modular and has set of services. Each service offers some task to do which then combines its data to give fully functional virtual assistant. Following is a brief idea about how the virtual assistant is going to function.

It starts with the first step of facial recognition. If the user is detected it transfers to the next step else the prompt is provided as "User not detected want to register as new user" and new user registration prompt is opened and the predefined quaternary is loaded and the user is asked to answer the following questions for the registration process. Once all the questions are answered the facial sample photo is collected and the user is registered successfully and the application starts from the beginning.

Once the user is detected the application is connected to the database having the data of the particular user and the assistant is ready for the query. The user can start the conversation ask a question or do as the user wish. The speech recognition program converts the speech of the user into the text format and saves that information into the user database as the future data for speech recognition. The generated text is then transferred to the Chatbot application or can be called as dialogue manager.

Then the proper reply is generated using the knowledge database. Once the reply is generated the text is then converted into speech and the output is produced through speakers.

IAVA is mainly divided into three services which handle the most of data. The following is the services we proposed in the IVA:

1.1 Face Detection Service

The Face Detection Service allows the virtual assistant to detect the presence of the user in front of the device and verify its user data using the face in the image and database. Face Detection Service continuously scans the video input from the camera or webcam. As soon as face is detected virtual is available for further query. Face Detection Service uses the Deep Learning method to detect the face and authorize the user.

1.2 Speech Detection Service

The Speech Detection Module allows the virtual assistant to record the user's voice data using the microphone which then stores into the user database for the speech recognition. It also has the functionality of speech synthesis which converts the text on screen to the audio.

1.3 Dialogue Manager Service

The Dialogue Manager is the soul of a virtual assistant as it generates the query reply using its knowledge database. It has the functionality to give the most effective and best reply to the query asked by the user. The user input is

mostly textual or vocal which the processed using the service which is used in the Dialogue Manager. Dialogue Manager is the key service which has the most complex task to do and give the accurate reply to the query.

1.4 Database

In this Virtual Assistant we divided the database into two part which is as follows:

1.4.1 User Database

The user database has all the information about the user which its image and vocal voice. It serves for user authentication and user insertion.

1.4.2 Knowledge Database

The knowledge database can be local as well as online which include the facts about the user and its queries and reply's database which gives the idea about how user and reply generation.

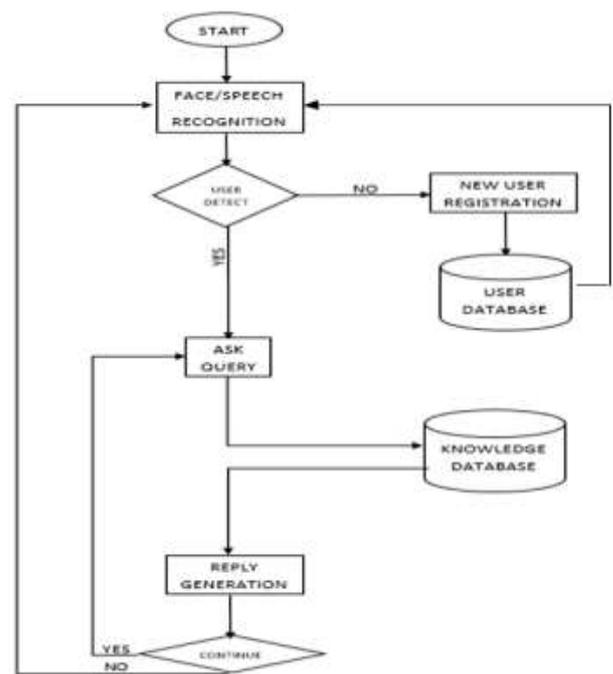


Fig-2: Flowchart

4. RELATED WORK

In the paper [1], the authors have explained how virtual private assistant work and how they are being upgraded using various new technology. It is the multimodal dialogue system. VPAs system has used speech, graphics, video, gestures and other modes for communication in both the input and output channel. Also, the VPAs system will be used to increase the interaction between users and computers by using some technologies such as gesture recognition, image/video recognition, speech recognition, and the Knowledge Base. Moreover, this system can enable

a lengthy conversation with users by using the vast dialogue knowledge base. Our project emphasizes on the VPA being device-independent which can be accessed whenever and wherever wanted.

In the paper [2], the authors have explained the AR-based Assistant which combines the human interface and location-aware digital system. It gives the much rich experience to the user. In this project they are closer to create the virtual personal manager which gives the idea about it surrounding and location using augmented reality.

In the paper [4], the authors have explained about smart assistants and smart home automation which gives the idea about speech-enabled virtual assistants which they find less secure so using different technique they tried to overcome that issues.

5. CONCLUSIONS

Our paper introduced IAVA – our Omni accessible virtual personal assistant which can be accessed from any device and can be used by any registered user. We propose to utilize various AI technique to achieve so such as face detection, speech recognition, Chatbot application, text to speech translation. All this while providing an interactive animation.

Based on our data we can find that this type of project can be very popular in user since it can be accessed from any device and it can be used in the future project. This type of Projects can be used for medical purpose, business purpose, and many other applications.

6. FUTURE SCOPE

Following technology further can be upgraded with new budding techs such as emotion detection and live face interaction. The interactive animation can be upgraded to the facial animation for more human-like feel.

ACKNOWLEDGMENT

We would like to show our gratitude to the professors for sharing their pearls of wisdom with us during the course of this research, and we thank them for insights and for their comments that greatly improved the report. We are also immensely grateful to our department for providing all the necessary equipment and facilities.

REFERENCES

- [1] *Nextgeneration of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa, and Google Home) by Veton Këpuska; Gamal Bohouta 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*
- [2] *MARA – A Mobile Augmented Reality-Based Virtual Assistant*
- [3] *G. Bohouta, V. Z. Këpuska, "Comparing Speech Recognition Systems (Microsoft API Google API And CMU Sphinx)", Int. Journal of Engineering Research and Application 2017, 2017.*
- [4] *A vision and speech-enabled, customizable, virtual assistant for smart environments*