VOX – A Desktop Voice Assistant

Aniket Murkutkar¹, Aradhya Bobade², Naina Kaushik³

1,2SE Student, Computer Engineering, Rajiv Gandhi College of Engineering, Andheri, Maharashtra, India
3Assistant Professor, Dept. of Computer Engineering, Rajiv Gandhi college of Engineering, Andheri, Maharashtra, India

---***---

Abstract - V*oice assistants are programs on digital devices that listen and respond to verbal commands. This voice assistant will gather the audio from the microphone and then convert that into text, later it is sent through an API . The API will then convert text into audio file in digital language(binary) which is then converted to English language. Voice assistant is a technology which allows us two basic operations that are speech to text conversion in which we will have the user give a voice command and after getting an input the assistant will display what the command user gave and then implement it and text to speech conversion in which we will have the user write a text command which the assistant will interpret and give us a speech translation of the written command.*

Keywords— *API, speech to text, text to speech, verbal commands, voice assistant.*

1. INTRODUCTION

Voice assistant is an application program that understands natural language and voice commands to complete tasks for the user. The users can even ask their assistant questions and also interact with them. Voice assistant is used to run machine like laptop or PC's or a mobile phone on your own voice command. For this software it uses a microphone as input device to receive voice requests from user and speaker as output device to give the output voice.

1.1. Advantages of using Voice Assistant

1. Hands-free operation: Users can interact with the assistant using their voice, eliminating the need to use a keyboard or mouse.

2. Multitasking and seamless integration: Desktop voice assistants often integrate with various applications and services, allowing users to seamlessly switch between tasks or access information from different sources.

3. Accessibility and inclusivity: Voice assistants make technology more accessible to individuals with disabilities or those who have difficulty using keyboard and mouse. People with mobility issues, visual impairments, or certain medical conditions can benefit from voice assistants.

4. Efficiency and productivity: Users can simply speak commands or ask questions, and the assistant will respond with the relevant information or perform the requested action.

5. Expanded functionality and integration: Increasing compatibility with various applications, platforms, and devices allows for a more seamless user experience. For example, integrating with popular productivity tools, smart home devices, or third-party services can provide users with a wider array of capabilities.

6. Customization and user control: Providing users with options to customize the voice assistant's behavior, voice, or interface can enhance the overall user experience. Allowing users to define preferences, choose different voices, or customize settings empowers them to tailor the assistant to their individual needs and preferences

1.2. Drawbacks of current system

1. Privacy concerns: Users may feel uneasy about having a device that is always listening to their conversations, as it may lead to unintended data collection or potential breaches of privacy.

2. Limited functionality and compatibility: Compared to their mobile counterparts, desktop voice assistants typically have limited functionality. They may not support certain features or integrations available on mobile devices, which can limit their usefulness.

3. Misinterpretation and errors: Variations in accents, speech patterns, or background noise can affect the accuracy of voice recognition.

4. Dependency on internet connectivity: Without a stable internet connection, the functionality of the voice assistant may be limited or even rendered unusable.

5. Difficulty in Multitasking: In a shared office environment or a quiet library, users may prefer traditional input methods to avoid disturbing others.

6. Limited Use Cases: Voice assistants have limitations in certain domains. They excel at tasks like setting reminders, answering general knowledge questions, or controlling smart home

devices. However, they may struggle with complex tasks that require deeper domain expertise or lack integration with specific applications.

2. SUGGESTED IMPROVEMENTS

• Language interpretation should be improved as Indian speakers have a thick accent. Adding support for additional languages allows users who are more comfortable communicating in their native language to interact with the assistant effectively.

• For the purpose privacy and information protection a system supporting personal voice commands on the desktop assistant.

• Offline functionality of Operating System commands will let us perform operations even without internet connection.

• An android based voice assistant can have increased range of operations for wide usage of voice assistant.

• Introducing multimodal interfaces that combine voice with visual and touch inputs can enhance the user experience. Adding support for screens or displays allows voice assistants to provide visual feedback, display information, or enable users to interact using traditional touch or mouse inputs when appropriate.

• Voice assistants can be improved by enlarging their knowledge base and domain expertise. Integrating with more external services, APIs, and databases can enable voice assistant to provide accurate and up-to-date information across a broader range of topics and domains.

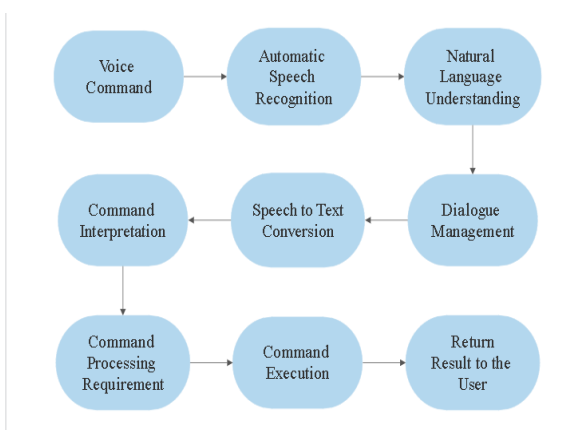## 3. DETAILS RELATED TO PROPOSED SYSTEM

Figure 1. Flowchart of system

1) Speech Recognition

In this module the system receives user commands and converts spoken words or phrases into text. The text is then transcribed into a form that can be understood by the system.

2) Natural Language Understanding

The text which is converted to a form understood by the system is analyzed to understand the user's intent and interpret the commands, queries and instructions given by the user.

3) Task Execution

This module is responsible for carrying out specific tasks or actions requested by the user. It interacts with other software components or services to perform functions such as setting reminders, sending emails, playing music, providing weather updates, or any other supported actions.

4) Speech Synthesis

The speech synthesis module, also known as text-to-speech (TTS). It converts the system-generated text responses into spoken words and transforms the text into an audible output that the user can hear through the speakers or headphones. It can also convert a text input into speech.

5) User Interface

The user interface module provides the visual representation of the voice assistant on the desktop screen. It includes elements like graphical user interface (GUI), settings, help option for user, command list for better user experience, clear screen, etc.

4. USER INTERFACE

• Main Page

GUI for the desktop voice assistant. When the speak button is activated, the system starts to take the user's voice commands.
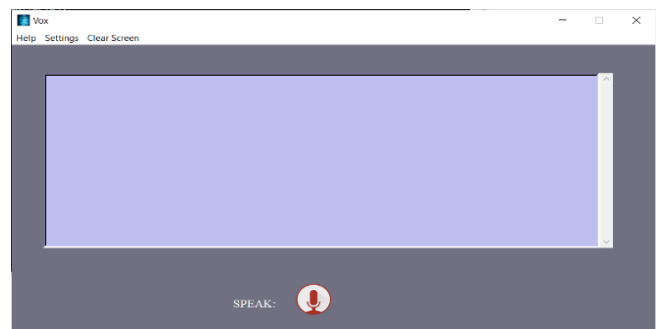

Figure 2. Main Page

- Welcome Page

The system greets the user, asks his/her name for future reference and signals the user to give the command.

Figure 3. Welcome Page

- Running Commands

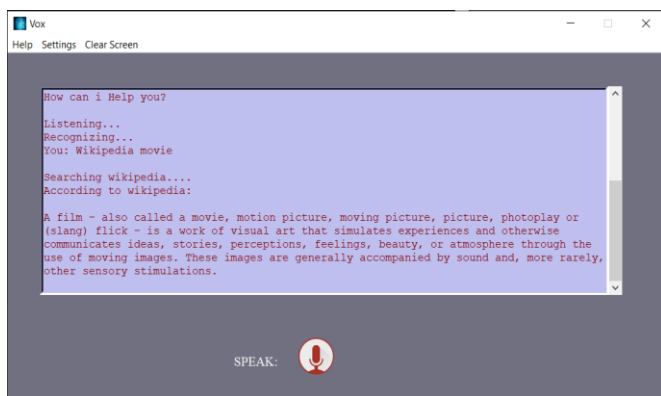After taking the user desired input, it executes the command.

Figure 4.1. Command Execution

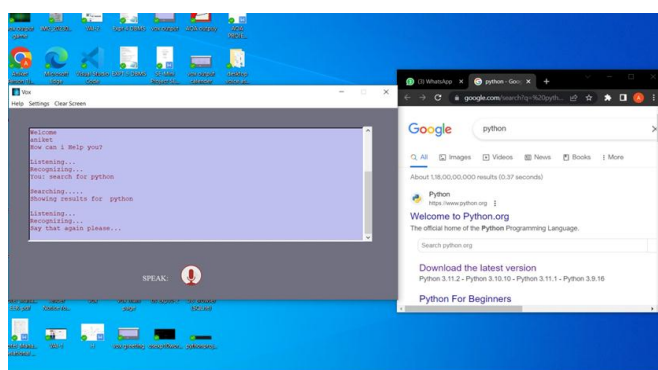In the above figure, the user says 'Wikipedia movie' which implies to describe the word movie through Wikipedia.

Figure 4.2. Command Execution

5.FUTURE SCOPE

As per the project perspective, it can be improved to an android application which can increase the mobility of the voice assistant from accessing it just from the laptop/PCs. Design work will include the improvement in the design of the GUI. For betterment of user, the AI bot can be upgraded into a personalized voice assistant. This means that the user will get to add their own voice by which only their voice will be recognized by the artificial intelligence and work only on the user's voice command. This will result an upgrade in the security aspect of the application.

6.CONCLUSION

Voice Assistant is the application of artificial intelligence and Software development where we are creating a voice assistant which helps us to remotely access our hardware through voice commands. This project will benefit all the people by saving the time to type and rather just speak and control the device. Voice assistant can prove to be beneficial on both industry and personal level for the users.

REFERENCES

[1] "Research Paper onDesktop Voice Assistant", International Journal of Research in Engineering and Science (IJRES), vol.no.10, page nos.15-20 , year 2022.

Vishal Kumar Dhanraj (076) Lokeshkriplani (403) Semal Mahajan (427), "Gradient-based learning applied to document recognition,‖ Proceedings of the IEEE, vol.10, pp.15-20, 2022.

[2] "Donna – A web based AI Personal Assistant", International Journal of Computer Applications, vol.175-no.8 , page nos.07-12 , year 2017. Ami Doshi, Ria Shah,Drasti Bhimani, Bhoomi Patel, Swati Mali, IJCRT, ISSN: 2320-2882

[3] "Desktop Assistant Theory Explained," https://www.geeksforgeeks.org/Desktop Assistant-theoryexplained/,

[4] "Desktop Voice Assistant", International Research Journal of Modernization in Engineering Technology and Science, Volume:02/Issue:05/May-2020