Holographic Smart Home Assistant

Aswath K[1], Albin Reji[2], Rahul K[3], Nirmal C[4]

[1234]U.G Scholar, Department of Computer Science and Engineering
Shree Devi Institute of Technology, Karnataka, India

Abstract - *Smart virtual assistants have been in the technology world for a while. Technology giants such as Google, Apple, Amazon introduced their own smart Artificial Intelligence virtual assistants which opened a new way of gathering information from the web, managing connected devices and software. The main goal of an AI assistant is to work for us and help us in making our life better.*

This paper primarily focuses on developing a smart virtual assistant inside a 3D holographic display. The system makes use of open source Artificial intelligence tools for creating the smart assistant. The system runs on top of a Raspberry Pi, which has the connectivity options such as Ethernet and WiFi which can be used to connect to the internet for knowledge-based queries and other tasks.

Key Words: **Artificial Intelligence, Virtual Assistant, Holographic Display, Speech Recognition, Pepper's Ghost Effect**

1.INTRODUCTION

Virtual assistants are interactive systems which act as real-world assistants that can perform tasks or services for an individual. With the latest advancements in Artificial Intelligence and Machine learning, the virtual assistants are capable of doing tasks which outperforms real-world assistants. With the improved data processing capabilities and availability of data, Latest deep learning algorithms are used for tasks such as Computer Vision and Natural Language Processing. Now the machines are able to understand the context in a sentence better than human beings.

The battle of virtual assistants began when the tech giants introduced their own virtual assistants such as Google Assistant, Apple's Siri, Amazon's Alexa, Microsoft's Cortana. Mostly they interact with the user via so provided Chat interfaces. The major issue with these assistants is the interface that we are using to interact with them. They are the typical chat UIs or voice only UI. We are not getting a visual feedback that attaches to the real world. We are receiving either the audio only or audio with some lights blinking.

In this paper, we will focus on developing a virtual assistant where the users will be able to interact with it via a holographic display and built-in microphone. The above said holographic display is constructed using an old technique called as Pepper's Ghost effect, which shows a fake-holographic 3D display. We will make use of open source AI software tools and libraries which are available for free of cost. The entire system will be running on top of Raspberry Pi.

2. SYSTEM ARCHITECTURE

The entire system runs on top of Raspberry Pi, a single board computer. This has the connectivity options such as Ethernet and WiFi, which can be used to connect to the internet. MIC is attached to the system so that the voice commands from the user can be recorded. A display and speaker is attached so that the users can be provided with a visual and audio feedback on their actions.

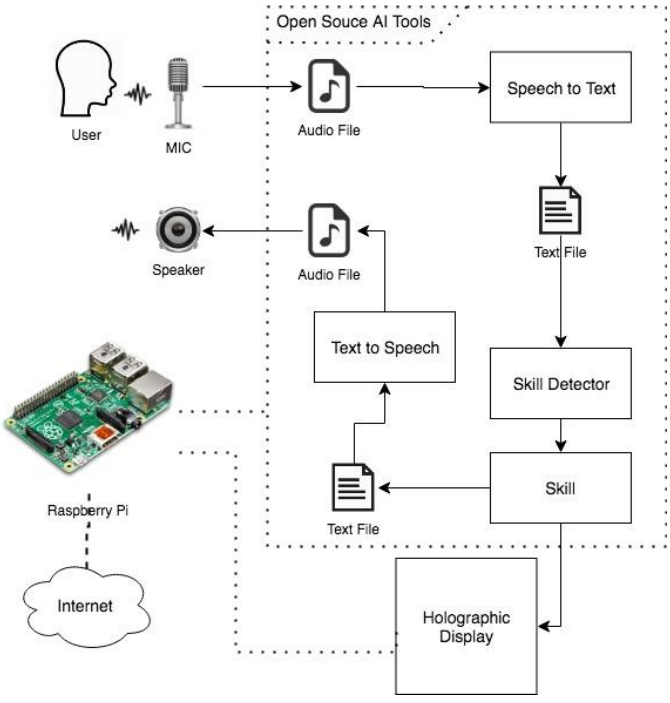

Fig -1: System Architecture

2.1 Holographic Display

The display is created based on a technique called as Pepper's Ghost effect. Even though we call it as a holographic display, it doesn't have the qualities of a real holographic display. It's a pseudo holographic display. Pepper's Ghost effect uses a simple technique which basically reflects an image off of a surface to create an

illusion of 3D object floating in space. 4 Glass pieces are arranged in such a way that maximum reflection from the display is reached into the user's eyes.

The monitor contains a headless browser window which has the side views of the 3D model placed in such a way that the glasses will reflect the contents of the monitor in all views. The browser window contains a Three.js program which is connected to the socket from the AI backend. With respect to the message received from the socket, the 3D character creates animations, which will get reflected in the glass arrangement.

2.2 Open Source AI Tools

For the software part of the system, we make use of open source tools and libraries such as gTTS, SpeechRecognition, pocketsphinx, adapt etc.

When the system is started, it listens for the pre-registered wake word. The system remains to do no other activity until wake word is detected. When it detects wake word, the system starts to record the audio preceding the wake word. The recorded audio is sent to the Speech to Text engine, which converts the recorded audio into corresponding human-readable text format. This text file will be sent to the skill detector module, which does the Natural language processing task and figures out the intent of the sentence that user spoke. The generated intent is compared with the installed skills. If any match found, that particular skill will be invoked. Otherwise, default skills such as Web search skill will be called. Next, it is the task of that skill to do the remaining task.

We use padatious, An efficient and agile neural network based intent parser to parse task the user intends to accomplish from the textual representation. The machine learning model in Padatious requires a relatively small amount of data. PocketSphinx is a lightweight speech recognition engine, specifically tuned for handheld and mobile devices, though it works equally well on the desktop. We will make use of WebSockets to facilitate the communication between modules in the system. When the intent is figured out, the specified skill will be called. The work will be done by the registered skill. The skill may make an audio output through the speakers or not.

3. CONCLUSION

This paper summarises the issue with the current state of virtual assistant and provides a workaround for that. We were able to create an affordable smart AI assistant with the help of Open Source tools. The system will be able to serve its user very well with the audio and visual interaction. Since the system is architectured properly, others can implement new features to the system just by adding new skills to the system. Anybody can create a new skill for the system using the documentation available online.

REFERENCES:

[1] Abhay Dekate, Chaitanya Kulkarni, Rohan Killedar, "Study of Voice Controlled Personal Assistant Device", IJCTT, Vol 42, No 1, Dec 2016.

[2] Sheela Deshmukh, Neha Jundare, Mohini Gore, Harsha Sarode, “Speech to Text Recognition System,” IRJET, vol. 03, Issue 10, Apr. 2017.

[3] Nuzhat Atiqua Nafis & Md.Safaet Hossain, “speech to text conversion in real time”, Retrieved on 2 Aug 2015.

[4] http://www.instructables.com/id/Reflective-Prism/