

Voice Controlled Multi-Functional Smart Device

Samidha Ajgaonkar¹, Mansi Chauhan², Hitesh Humane³, Kartik Jain⁴, Ankita Malhotra⁵

^{1,2,3,4} UG Student, Department of Electronics and Telecommunications Engineering, Rajiv Gandhi Institute of Technology, Maharashtra, India

⁵Assistant Professor, Department of Electronics and Telecommunications Engineering, Rajiv Gandhi Institute of Technology, Maharashtra, India

Abstract - This paper proposes a novel device capable of performing multifunctional operations allowing the user to perform various tasks simultaneously like controlling home appliances, browse any information or knowing about any incoming mail. This system provides interaction between users and devices using web applications. The device can access all web text with the help of voice along with incorporating a smart home automation system which is also voice-enabled along with a voice notification system for emails. Speech recognition system built in Python is used. BeagleBone Black has been used as a central controlling processor which will convert the input voice commands into machine-readable inputs, thereby making the machine ready to perform the tasks requested by the user. Thus, a voice-controlled smart home automation system is designed, so that the users can perform certain tasks by just the use of their voices.

Key Words: Automation, BeagleBone Black, Python Voice-controlled system, Voice Notification.

1. INTRODUCTION

Devices using Voice user interface has been a hot topic for the past few years. With the rise of Alexa and Google Assistant among others, people are growing increasingly open to this new technology [1]. Voice-enabled devices are the primary or supplementary visual, auditory, and tactile interfaces that enable voice interaction between people and devices, using speech recognition to understand spoken commands and text to speech to play a reply. Simply stated, a Voice-controlled smart device can be used for anything from a light that blinks when it hears your voice to an automobile's entertainment console [2]. The motivation behind such a project is that it is highly useful for performing multiple operations simultaneously thereby reducing the user's effort and time. It is particularly helpful for people with disabilities or for old age people as through it they can control the various appliances around the house, browse any information in real-time by just giving voice commands, or receive voice notification for any incoming emails [3]. Earlier work done in this field proposed a Bluetooth Based Home Automation System where the system used Arduino and relay for controlling home appliances or a Zigbee based wireless home automation system controlled by voice commands [4][5] [6] [7]. In this model Wi-fi has been used for wireless connectivity which is advantageous over other networks like Zigbee or Bluetooth in terms of convenience,

mobility, speed, and range. Another design methodology proposed a Digital Assistant for the Blind which enabled the users to receive and send emails, access daily news, weather forecasts, etc. through voice [8]. A Voice Controlled E-Commerce Web Application used IBM Watson to present a voice-enabled commodity purchase e-commerce application [9].

In this work, a multi-functional device is proposed which can perform multiple tasks like home-automation, voice search, and notification in a single system. Another paper proposed to provide an easier interface between the user and web application where the entire web text was accessed with the help of voice using JSAPI components from the Java [10]. However, in the present work, open-source libraries like Pocket Sphinx and Pytttsx respectively available in Python, have been used for Speech To Text (STT) and Text To Speech (TTS). This paper proposes the implementation of a voice assistant on a BeagleBone Black, using a USB microphone for voice input and a 3.5 mm audio output speaker. This is an advantage in comparison with [11] where Raspberry Pie is used as a central controller which is relatively slow for multifunctional operations.

The rest of the paper is organized as follows: Section II presents the proposed system architecture showing hardware and software requirements along with a block diagram. Section III presents the operational flow graph of the device along with the methodology. The actual working proof is presented in section IV. Section V concludes the paper with a discussion of future work.

2. PROPOSED SYSTEM ARCHITECTURE

Beaglebone Black acts as a central controlling unit taking user input through a microphone and giving voice output through a speaker. Tactile switches and LED's make-up for the user interface system. For wireless connectivity wi-fi module has been connected to BeagleBone and relay module. Depending on the user input the actions are performed, either for voice search or controlling appliances.

The voice enabled system begins its functioning by taking the user input through the microphone. The next step is Speech Recognition where audio is converted to words. The next step corresponds to Understanding the spoken language where words are converted to meaning. Then follows a dialogue manager who decides the next action of the system. Following the dialogue manager comes the response generating segment where meaning is converted to

words. Text to speech synthesis is done where words are converted to audio and an appropriate system's output is given. All of the above steps are followed for both the home automation and voice search segments wherein the response depends on the user input. For email notification, no user input is required. The device acts as a simplex device in this case.

The design for the proposed model is shown in Figure 1.

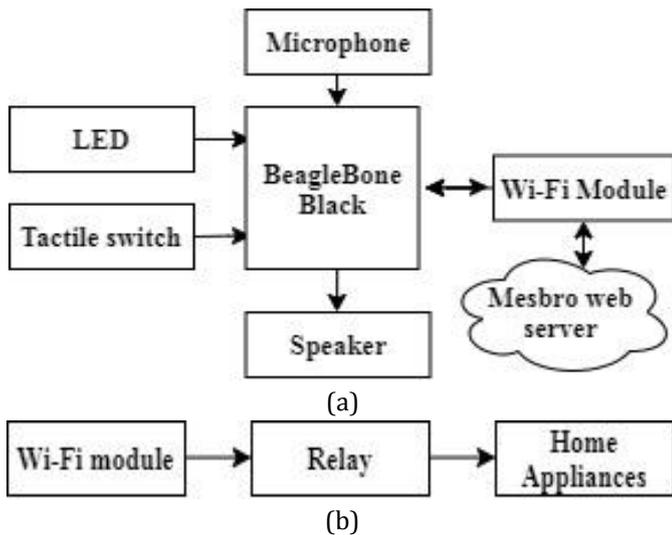


Fig. 1 Block diagram (a) BeagleBone interfacing with user interactive devices (b) Wireless control of home appliances

For home automation, we speak the predefined commands to the microphone connected to BeagleBone Black using the python speech recognition system. The commands are then sent through Wi-Fi between BeagleBone and relay and perform the desired task. At the same time, the commands are displayed on the screen. Each command defined in code have a distinctive operation. We can change the commands according to our ease.

The device has an email notification system allows for more efficiency when it comes to communicating through email. The said system is an application that notifies users when email message comes in. This device notifies users of new messages even if the user is unable to check its inbox or if the mailbox is not open. Notifications can further be customized according to the format that works best for the user. It notifies without disrupting the user if they are working on a computer or can be muted as and when needed. The notification system has been incorporated using the Mesbro mail API.

A. Hardware components

- **BeagleBone Black:** It is a development board based on the AM3358 processor. It has 4GB memory in which the operating system can be loaded and the coding can be done using cloud9 IDE.
- **Wi-Fi module:** The module used in this project is ESP8266-01. It has an integrated TCP/IP protocol stack that

gives the BeagleBone Black development board access to the internet.

- **Relay module:** It is a digital switch that is used for switching voltages and currents. The relay performs switching actions based on input provided by the user.
- **Microphone:** It is a sensor that detects voice queries and sends electrical signals for further processing.
- **Speaker:** It is an actuator that converts the delivered query from the processor into voice output.

B. Software tools

- **Cloud9 IDE:** It is a cloud-hosted toolkit. For BBB, it provides an integrated, open-source development environment to build projects. It is also very flexible with different programming languages. In this model cloud9 has been used cloud9 for system programming using python and also used to establish a web server connection with BeagleBone black.
- **Natural language processing toolkit:** NLTK is a platform for building a python program that works with human language data. It consists of various text processing libraries for classification, stemming, tokenization, tagging, parsing, and semantic reasoning. It is a free, open-source, community-driven project used for categorizing text, analyzing linguistic structure.
- **Mesbro web server:** It is used to access the user's personal account for email notification purposes.

The Breadboard view of the proposed model is shown in figure 2:

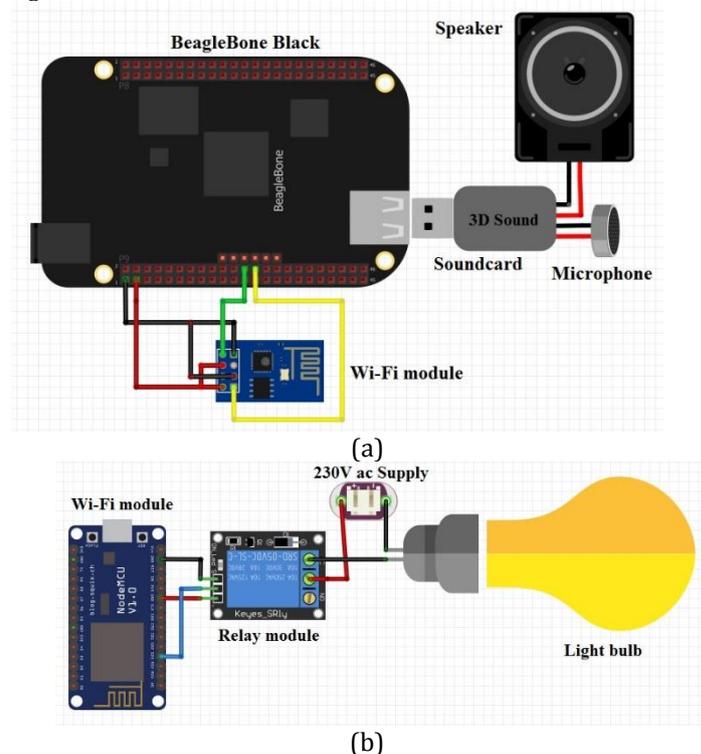


Fig. 2 Breadboard view (a) BeagleBone interfacing with user interactive devices (b) Wireless control of home appliances

3. PROPOSED METHODOLOGY

- Initialization: As shown in figure 3a, Once the device is turned 'ON', it checks for internet connectivity. If proper internet access is available then the user authentication process is initiated. If the credentials are found to be correct the device will give a voice acknowledgment. This will let the user know that the device is powered ON. The user can then send voice queries which the system will recognize using a microphone. The voice input is processed and transcribed into text. The text is analyzed to detect questions and commands. Depending on the input keyword the desired functionality is performed.
- Voice notification: As shown in figure 3d, for voice control, email notification system the device is required to check the user database within a stipulated time. If the user is having a "new mail" in the inbox, the user is notified through voice command ("YOU HAVE RECEIVED A NEW MAIL").

The working of various segments of voice control smart device is shown in figure 3:

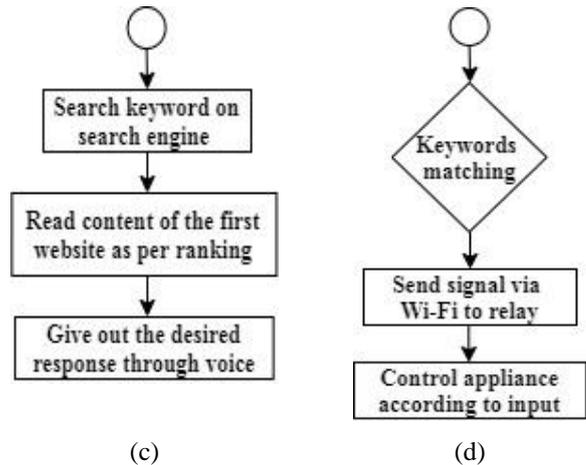
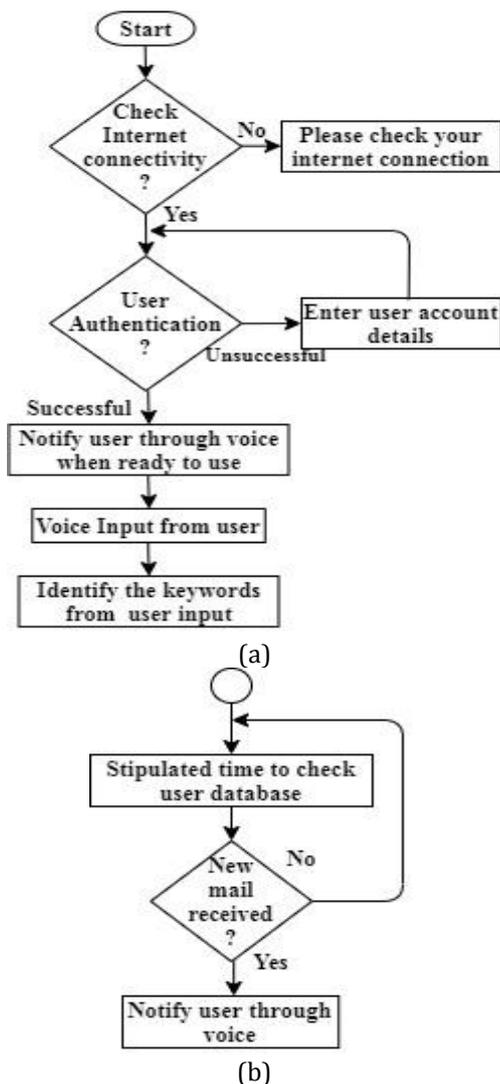
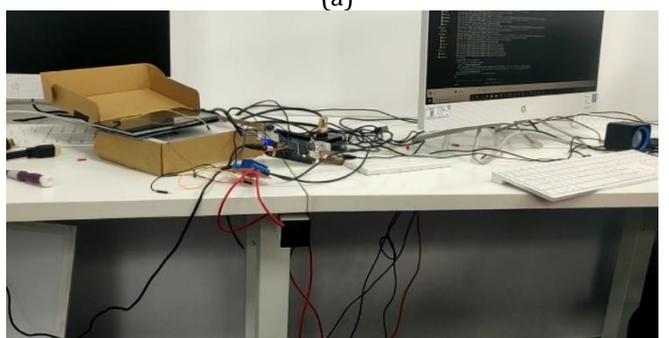
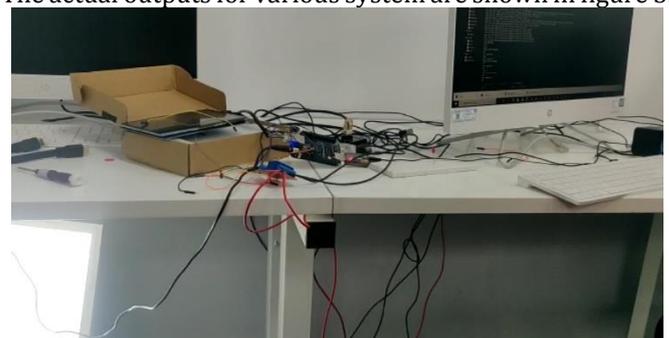


Fig. 3 Operational flow graphs (a) Device initialization, user authentication and input (b) Email Notification (c) Voice Search (d) Automation

- Voice search: As shown in figure 3b, Voice search system analysis the text and connects to external data sources such as a search engine to find the relevant information. The information is translated into a digestible format and fulfills the user's intent.
- Home automation: As shown in figure 3c, A home automation system makes use of predefined keywords for controlling home appliances. If the input from the user matches the keyword (e.g., "LIGHTS ON"), the processor will send a signal to relay via Wi-Fi and control the appliance.

4. RESULT

The actual outputs for various system are shown in figure 3.



```

root@beaglebone:~# python3 Speech.py
say something:
turn on light
Light ON
You have received new email
say something:
virat kohli
Virat Kohli is an Indian cricketer who currently captains the India national tea
n. A right-handed top-order batsman, Kohli is regarded as one of the best batsme
n in the world.
say something:
turn off light
Light OFF
You have received new email
say something:
    
```

(c)

Fig. 4 Actual working demonstration of the system

(a) Switching lights ON (b) Switching lights OFF (c) Output for complete system

Figure 4a and 4b shows the output for voice-controlled switching “ON” and “OFF” of light. The set keywords for performing the control activity are “light ON” and “light OFF” respectively. Once the user pronounces the keywords the complete process for achieving the desired operation is initialized.

Figure 4c shows the output for the email notification system along with voice search. Upon receiving new mail, the user is notified vocally. For voice search, the response is dependent on what information the user wants to browse. For all the three systems the output is printed on console as shown in figure 4c.

Commands: Following commands should be spoken by the user to turn on and turn off devices.

Table 1 Commands for controlling home appliances

Command send by user	Message display on monitor
turn on light	Light ON
turn off light	Light OFF
turn on fan	Fan ON
turn off fan	Fan OFF

5. CONCLUSIONS

A voice-interactive IoT device has been presented to improve accessibility to web applications. This model aims to enhance usability for all users by promoting ease of interaction, multitasking and support a lean environment where users can make requests using natural language. The speech recognizer that we used in our proposed system for making better interaction of web applications with the help of voice is open source, developed by “CMUSphinx”. This recognizer is capable of live speech, medium vocabulary speech recognition. The implemented system is developed using BeagleBone Black as the central processing unit, ESP8266 for providing wireless connectivity and relay as a

switch. Further, the proposed model can be extended to incorporate calling facility, multi-language support, and improved voice search result. A lot of work is also been done to incorporate artificial intelligence technology into this system.

6. ACKNOWLEDGEMENT

Authors thank Mesbro Technologies Pvt. Ltd. for financial and facilities support to this research publication. Also thank our colleagues in Rajiv Gandhi Institute of technology for supporting our team.

REFERENCES

[1] Rebeca Costa, Voice user interface design: a guide for beginners, JUSTINMIND, Sep. 13, 2019. Accessed on: Nov. 3, 2019. [Online]. Available: <https://www.justinmind.com/blog/voice-user-interface/>

[2] Justin Baker, Voice User Interface (VUI)- The Ultimate Designer’s Guide, MUZLI, Nov 25, 2018. Accessed on: Oct. 4, 2019. [Online]. Available: <https://medium.muz.li/voice-user-interfaces-vui-the-ultimate-designers-guide-8756cb2578a1>

[3] H. V. Bhatnagar, P. Kumar, S. Rawat and T. Choudhury, “Implementation model of Wi-Fi based Smart Home System,” International Conference on Advances in Computing and Communication Engineering (ICACCE), Paris, 2018, pp. 23-28.

[4] R. Piyare and M. Tazil, “Bluetooth based home automation system using cell phone,” 2011 IEEE 15th International Symposium on Consumer Electronics (ISCE), Singapore, 2011, pp. 192-195.

[5] A. Olteanu, G. Oprina, N. Tapus and S. Zeisberg, “Enabling Mobile Devices for Home Automation Using ZigBee,” 19th International Conference on Control Systems and Computer Science, Bucharest, 2013, pp. 189-195.

[6] A. J. Jara, “Wearable Internet: Powering Personal Devices with the Internet of Things Capabilities,” 2014 International Conference on Identification, Information and Knowledge in the Internet of Things, Beijing, 2014, pp. 7-7

[7] Y. Kung, S. Liou, G. Qiu, B. Zu, Z. Wang and G. Jong, “Home monitoring system based internet of things,” 2018 IEEE International Conference on Applied System Invention (ICASI), Chiba, 2018, pp. 325-327.

[8] P. Bose, A. Malphak, U Bansal and A. Harsola, “Digital assistant for the blind,” 2017 2nd International Conference for Convergence in Technology (I2CT), 2017.

[9] M. S. Kandhari, F. Zulkernine and H. Isah, "A Voice Controlled E-Commerce Web Application," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2018, pp. 118-124.

[10] D. Kumar and A. Sachan, "Bridging the Gap between Disabled People and New Technology in Interactive Web Application with the Help of Voice" IEEE International Conference on Advances in Engineering & Technology Research (ICAETR - 2014), August 01-02, 2014, Dr. Virendra Swarup Group of Institutions, Unnao, India.

[11] S. Mischie, L. M. Iovan and G. Gasprenc, "Implementation of Google Assistant on Raspberry Pi," 2018 International Symposium on Electronics and Telecommunications (ISETC), 2018.