User Interface for Visually Impaired People

D Kiran kumar

School of Electronics engineering VIT University, Chennai campus

Abstract: *The research work deals with the design and implementation of Speech recognition device using Raspberry Pi for essential usage to visually impaired people. Now a day's blind people are increasing enormously so, the main objective of the research is to provide a simple, cheap, friendly user, and compact device for visually impaired people to use multimedia applications of operating system like text, music player and dialing system by interfacing GSM module. All the above are implemented in a low cost Raspberry Pi board. Thus a goal of producing an compact device has been designed at low cost using offline speech recognition.*

***Keywords:*** *raspberry pi, sound card, GSM module*

I.   Introduction

The standard behind the model depends on speech technology. There are two sorts of speech technology. Speech synthesis is PC produced simulation of human speech. It is utilized to make an interpretation of composed data into aural data where it is more helpful, particularly for portable applications, for example, voice-empowered email and bound together informing. Blended speech can be made by connecting bits of recorded speech to make complete sentence. Speech recognition is a method to dissect the speech talked by a man that how it is vary to different personas with any way to deal with voice acknowledgment, the initial step is for the user to talk a word or expression into a receiver. The electrical sign from the amplifier is digitized by a "analog to digital (A/D) converter", and is put away in memory.

To decide the "signifying" of this voice input, the PC endeavours to coordinate the data with a digitized voice test, or format that has a known importance. This permits the user to control any gadget by voice, as opposed to using the mouse and console. The inserted framework which is changing over the speech information into content and by utilizing Google speech API or different applications like pocket sphinx the voice pursuit will be finished. The yield of voice pursuit will be put away into database and that sought information will be interface with solenoid plates by method for GPIO pins of raspberry pi. For this reason, the fundamental undertaking is to make GUI which depends on voice seek. To fulfil this reason, speech program is made by utilizing Embedded/QT maker instrument which is accessible in raspberry pi device chain. This speech program contains record catch, by squeezing the catch any individual can input a speech and the outcome taking into account voice hunt will be appeared in the window same as like Google pursuit. To fulfil this reason Google speech API is utilized. What's more, the consequence of hunt will be sent to GPIO pins. What's more, that sticks are associated with different solenoid plates. In this proposed work we interface the GSM to raspberry pi for dialling framework. That is helpful for the outwardly weakened individuals to impart to external world like ordinary individuals.

II.   System Design

A) *Hardware selection*

RASPBERRY PI: The Raspberry Pi is of charge card estimated single board created in UK discharged in 2012. It's elements are Broadcom framework on chip with ARM processor, speed extent is 700 MHz single-center ARM11 Broadcom BCM2835 CPU and Video center IV GPU, 512MB RAM, 2 USB spaces, HDMI port, Ethernet port, 3.5mm sound jack, keep running's with 5V small scale USB power connector and does not bolster settled capacity. SD card are utilized to store the working framework and project memory in either the SDHC or MicroSDHC sizes. It bolsters variety of Debian-based OSs, principally Raspbian OS. It ought to designed to Wi-Fi dongle. Most recent wheezy picture will bolster spilling and these picture is made uniquely to stream of information.
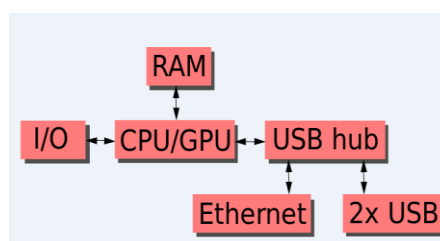
Raspberry Pi2 model B: The Raspberry Pi2 is additionally of Visa measured single board created in UK discharged in 2015. It's elements twofold the sum what was there in Raspberry Pi model B are Broadcom framework on chip with ARM processor, speed reach is 900MHz quad-center ARM Cortex-A7 Broadcom BCM2836 CPU and Video center IV GPU, 1GB RAM, 4 USB openings, HDMI port, Ethernet port, 3.5mm sound jack, keep running's with 5V small scale USB power connector and does not bolster settled capacity. SD card are utilized to store the working framework and system memory in either the SDHC or MicroSDHC sizes. It underpins Variety of Debian-based OSs, essentially Raspbian OS, free Windows 10 form. It comes as another Broadcom chip with four centers and a higher clock speed another point of preference accompanies the measure of RAM. Adding remote presentation elements to it bolsters as miracast. It ought to arranged to Wi-Fi dongle. Most recent wheezy picture will bolster gushing and these picture is made particularly to stream of data.

OPERATING SYSTEM	Same as for RaspberryPi1plusWindows10 additional variants Linux as Ubuntu and Android
CPU	900 MHz quad core ARM Cortex-A7
MEMORY	1 GB RAM
STORAGE	Micro SDHC slot
GRAPHICS	Broadcom Video Core I V
POWER	4.5 to 5.5 W

TABLE (1): Raspberry pi specifications.

GSM MODULE: The SIM900 is a finished Quad-band GSM/GPRS arrangement in a SMT module which can be installed in the user applications. Including an industry-standard interface, the SIM900 conveys GSM/GPRS 850/900/1800/1900MHz execution for voice, SMS, Data, and Fax in a little shape variable and with low power utilization. With a modest arrangement of 24mm x 24mm x 3 mm, SIM900 can fit all the space necessities in your M2M application, particularly for thin and minimal interest of outline.

USB MICROPHONE: It depicts adjustments to basic shabby outer USB associated soundcards to make them suitable for use with oscilloscope and range investigation programming for sound transfer speed applications. The least expensive gadgets seem, by all accounts, to be founded on the CM108 or CM119 controller with 16-bit ADC (receiver) information and 16-bit DAC (earphone) stereo yield. The recurrence scope of the MIC info can be produced using DC to around 20kHz (identified with 44kHz USB inspecting rate), in spite of the fact that the least difficult mod depicted here remains AC coupled thus stretches out down to around 5-20Hz. Programming known not valuable are Room EQ Wizard - REW 5.1, Virtual Analyzer, and TrueRTA (basic and simple to utilize yet costs). Programming like REW is intense and free, however can take some learning and does exclude an oscilloscope capacity.

MIC attachment info is mono, with ghost MIC controlling by means of 4.5VDC to attachment from chip inclination supply, and after that 100nF coupling top to CM108 MICIN pin. Alterations made: - uproot 1.2k smt interfacing with MIC attachment tip/ring, to expel apparition controlling from tip. - cut off ring terminal in the middle of attachment and pcb, so just tip and construct associated in light of MIC attachment. - add a 2M2 from tip to 0V to continue coupling top grounded. - include 680k in arrangement between ADC "MICIN" coupling top to MIC attachment tip. The MIC attachment information impedance is expanded to ~900kω. Viable ADC information resistance dictated by firstly checking when section begins with direct data (2.8Vpp ADC data, 1.0Vrms), with MIC recording level set to 0% to stop any increase settings. At that point add a resistor in arrangement to the MIC info to give an attenuator – with 680kω included, the new information level for section is ~4.2Vrms, so the divider proportion is ~4.2 : 1, and subsequently the successful ADC data resistance is ~210kω. Earphone attachment has 2.2VDC predisposition in light of the fact that decoupling tops weren't introduced (0R smt connections were utilized by producer). Changes evacuated 0R joins for Left and right yields, and included 33uF 10V coupling top for Right yield. So just right (tip) line yield associated with earphone attachment. The soundcard is not protected, so find far from switch mode flouro bulbs and mains power LED lights.

USB WIFI ADAPTER: TP-LINK's 150Mbps wireless and Nano USB Wi-Fi adapter allows users to connect a desktop or computer to a wireless network at 150Mbps. This mini adapter is designed to be as convenient as possible and once connected to a computer's USB port, can be left there, whether traveling or at home. It also features advanced wireless encryption and easy installation. It's main feature is to turn a wired internet connection from a router to a PC or Laptop by acting as a Wi-Fi hotspot. It's high performance leads to smooth video streaming, voice streaming and online gaming. The Network security Supports 64/128 WEP,

WPA/WPA2, WPA-PSK/WPA2-PSK security.

B) Software selection

CMUSPHIX: Voice innovations have quite recently been put into business utilize as of late in light of the fact that the registering asset required by SR (Speech Recognition) and TTS (text-to-speech) is substantial .In the mid-to late 1990s, PCs began to wind up sufficiently capable to bolster ASR/TTS .The two key hidden advances behind these advances are SR and TTS. SR intends to change our speech into content mode, TTS intends to change content into voice yield. The Sphinx is a speech acknowledgment motor framework created at CMU .It can be utilized to fabricate both little, medium and substantial vocabulary applications .Sphinx has been submitted to Source Forge as an open source extend, its source code is accessible for download structure the web.

ARCHITECTURE OF CMU SPHINX: Sphinx2 consists of a set of libraries that include core speech recognition functions as well as auxiliary ones such as low-level audio capture. These libraries are written in C Language. Its main features includes:

- Continuous speech decoding (as opposed to isolated word recognition)
- Speaker-independent (doesn't require the user to train the system)
- Ability to provide a single best or several alternative recognitions
- Semi-continuous acoustic models
- Bi-gram or tri-gram language models

A few components particularly planned for growing genuine applications have been incorporated into Sphinx2. For instance, numerous parts of the decoder can be reconfigured at run time. New dialect models can be stacked or exchanged progressively. Additionally, new words and elocutions can be included. The sound information can be naturally logged to petitions for any future investigation.

COMPONENTS OF CMU SPHINX: The abnormal state design for sphinx is direct. As appeared in the Figure 2, the engineering comprises of the front end, the decoder, a learning base, and the application the front end is in charge of get-together, commenting on, and preparing the info information. Furthermore, the front end extricates highlights from the info information to be perused by the decoder. The explanations gave by the front end incorporate the starting and completion of an information section. The learning base gives the data the decoder needs to carry out its employment. This data incorporates the acoustic model and the dialect model. The learning base can likewise get criticism from the decoder, allowing the information base to progressively adjust itself based upon progressive indexed lists.
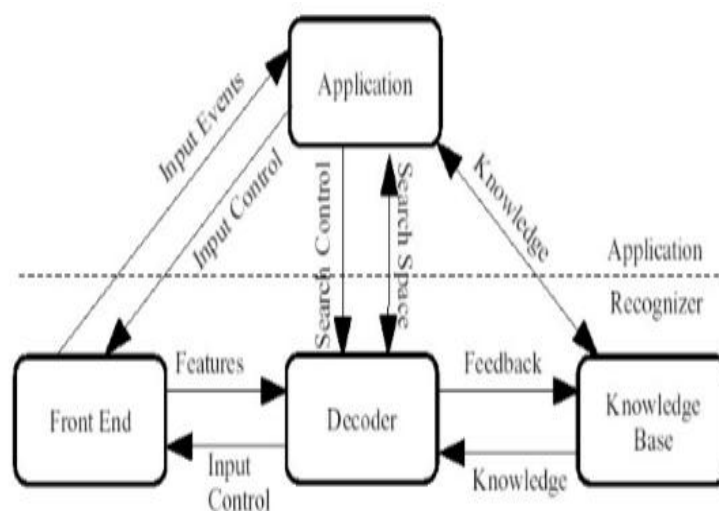

Fig (1): components of CMU SPHIX.

The decoder performs primary segment of SR. It chooses next arrangement of likely states , scores approaching components against these states , drop low scoring states lastly creates results. Points of interest of every parts will be talked about in the accompanying areas.

FESTIVAL: This application changes over the text to speech. Its simple include the system simply yield a string to the speech capacity rather than the screen. We needn't bother with a costly/muddled LCD or screen for our venture - simply utilize any old mp3 player amplifier or PC amplifier which you have most likely got lying around - or even a headphone functions admirably to debug purposes as well.

III.   Block Diagram

The basic block diagram of a voice recognition system is shown.the block diagram contains the speech recognition module i.e USB microphone for voice input because there is no input audio jac in raspberry pi, V power supply unit, limitted switches for manual operation,playback unit for voice output and interfacing of GSM module through GPIO pins of raspberry pi and external power supply for GSM module.
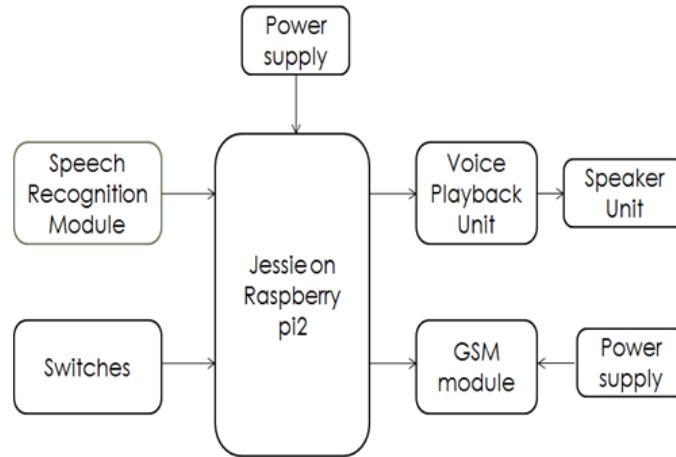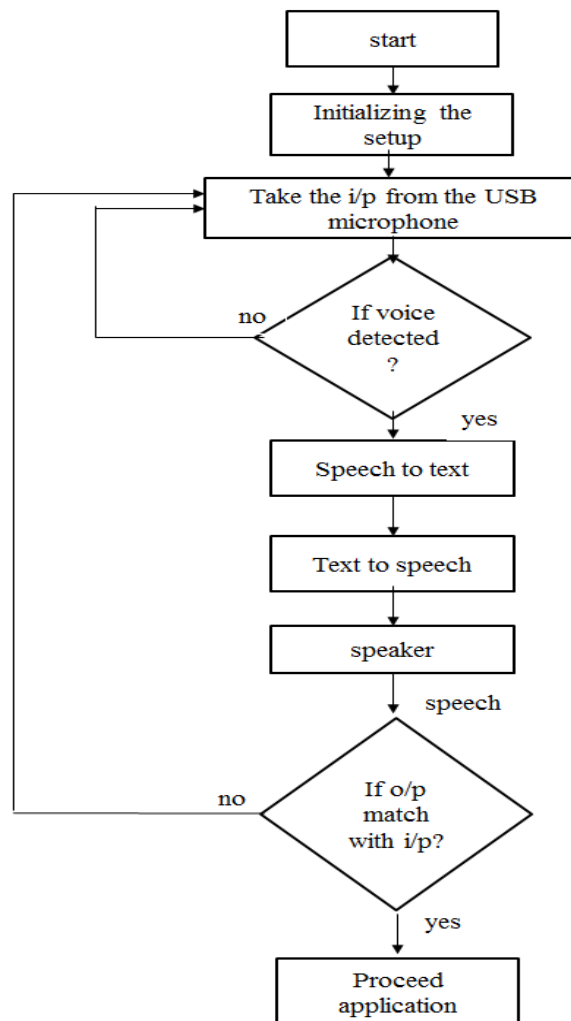

Fig (1): Block diagram of voice recognition system.

IV.   Flow chart

Fig (2): flowchart of voice recognition system.

V.   Implementation
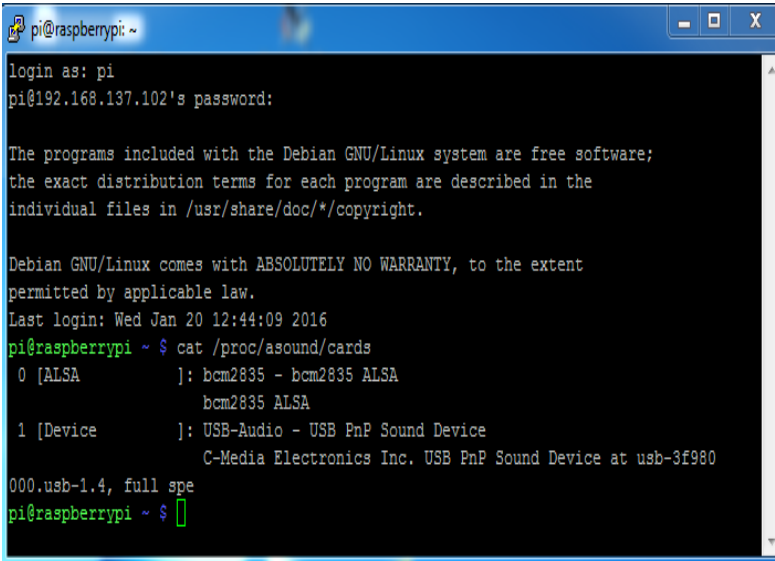
These can be achieved by configuring:
After choosing raspberry pi as hardware component we need to write a particular image (operating system which comfortable for our application).and setup the internet connection through wi-fi adapter which we chosen before.by executing the commands we need to update and upgrade the raspberry pi before installing the drivers related to our application.

Hardware setup:
a)   Install USB microphone related drivers.
b)   To check the USB microphone working properly or not by executing particular commands.
c)   Interfacing of GSM module to raspberry pi.

Software setup:
a)   Installing python
b)   Installing and run the pocket sphinx to raspberry pi by   executing particular commands.
c)   Installing and run the festival by executing particular commands.

Fig (1): USB microphone settings.

After installing both speech recognition application (pocket sphinx) and text to speech application (festival) we set up the USB microphone for taking the input voice. As shown in the flow chart.in this application we need to prepare the grammar file which contains the predefined words. We can upload the grammar file before or while executing we need to give the grammar file path. As shown in the flow chart the given input voice is converted to text by the application and that text format is input for the festival (text to speech application) through the speaker we can here the particular speech, if the input voice is matched to coming output speech that particular application is proceed if not, again it will take input voice from the USB microphone. This is the main objective of our project. For visually impaired people can use the compact device for multimedia applications of operating system like message reader, music player, and dialling etc.…

In this proposed work we implemented a compact device that includes a speech recognition system for dialling application. The main reason for design this particular system visually impaired people are facing difficulties to communicate with people through mobile phone. Voice is only option for visually impaired people to communicate with outer world. By using the speech recognition application system (pocket sphinx) we train the numbers from zero to nine or another option is we store the particular phone numbers with names in the raspberry pi data base. Before this we need to interface the GSM to raspberry pi. And we need to implement the two manual switches.one for power up the raspberry pi and the GSM module. And another for weather our voice (contained phone number or name) is desired or not. If the input voice (phone number) is match with output voice (phone number) we need to press that particular switch to proceed further action.
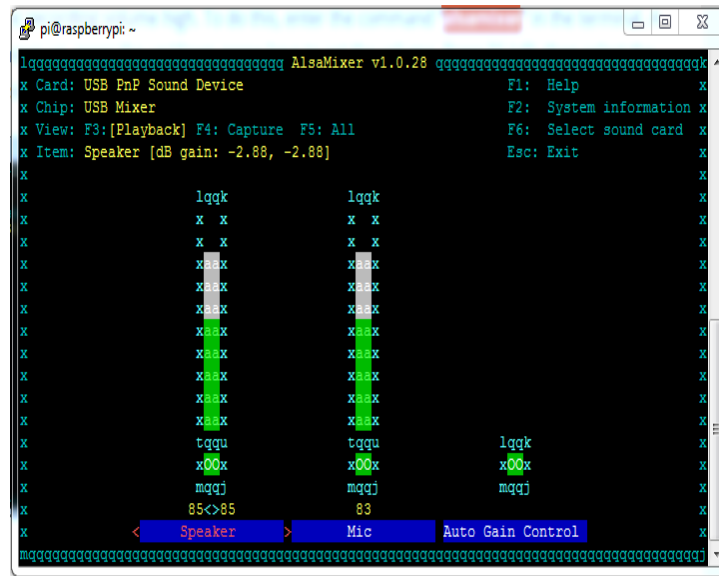
VI.   Results & Discussions

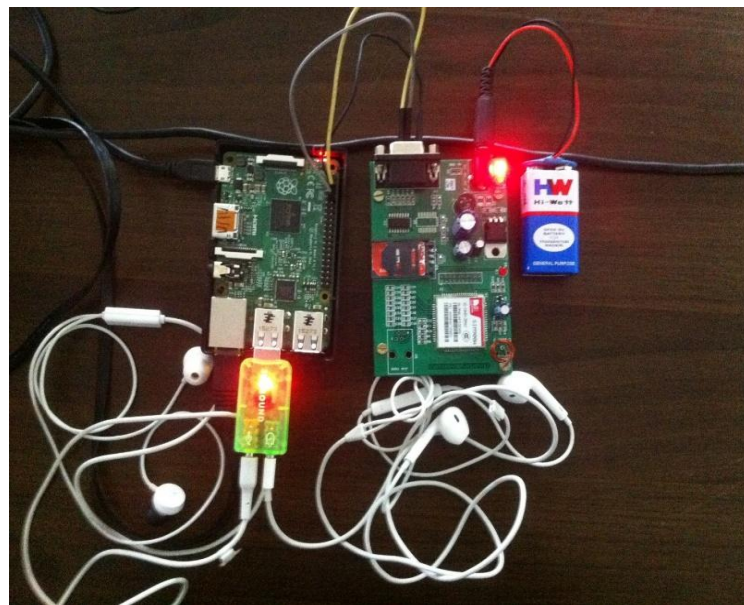**Fig (1):** speaker and microphone settings.

Fig (1): Block diagram of voice recognition system.

VII.   Conclusion

Visually impaired people experience difficulty and inconvenience using computers through a keyboard and mouse. The need to design this application is specifically for visually impaired people. Objective is to design an intelligent system based on speech recognition which will work according to voice search. by using this compact device visually impaired people can also communicate with outer world like normal people.

References

[1]. Sharma, Rupam Kumar, et al. "Android interface based GSM home security system." Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on. IEEE, 2014.
[2]. De Luca, Gabriele, et al. "The use of NFC and Android technologies to enable a KNX-based smart home." Software, Telecommunications and Computer Networks (SoftCOM), 2013 21st International Conference on. IEEE, 2013.
[3]. Gu, Yietal. "Design and Implementation of UPnP-Based Surveillance Camera System for Home Security." Information Science and Application (ICISA), 2013International Conference on. IEEE, 2013.
[4]. VanThanh Trung, Bui, and Nguyen Van Cuong. "Monitoring and controlling devices system by GPRS onFPGA platform." Advanced Technologies for Communications (ATC), 2013 International Conference on. IEEE, 2013.

[5]. Karia, Deepak, et al. "Performance analysis of Zig Bee based Load Control and power monitoring system." Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on. IEEE, 2013.

[6]. V. Tiwari, "MFCC and its applications in speaker recognition," International Journal on Emerging Technologies, vol. 1, no. 1, pp. 19-22, 2010.

[7]. T. Hain, P. C. Woodland, T. R. Nielser, and E. W. D. Whittaker, "The 1998 HTK System for Transcription of Conversational Telephone speech" in Proc. IEEEConf. on Acoustics, Speech and Signal Processing, Mar 1999, pp. 57-60.

[8]. J. G. Wilpon, "Automatic recognition of keywords in unconstrained speech using hidden markov model," IEEE Transactions on Acoustic, Speech and Signal Processing, vol. 38, no. 11, pp. 1870-1878, Nov. 1990.

[9]. A. G. Veeravalli, W. D. Pan, R. Adhami, and P. G. Cox, "A tutorial on using hidden markov models for phoneme recognition," in Proc. IEEE SSST'05, Mar. 2005, pp. 154-157.

[10]. T. Dasgupta and A. Basu. A speech enabled indian language text to braille transliteration system. In Information and Communication Technologies and Development (ICTD), 2009 International Conference on, pages 201–211. IEEE, 2009

[11]. R. Ghose, T. Dasgupta, and A. Basu. Architecture of a web browser for visually handicapped people. In Students' Technology Symposium (TechSym), 2010 IEEE, pages 325 –329, april 2010

[12]. T. Lauwers, D. Dewey, N. Kalra, T. Stepleton, and M.B. Dias. Iterative design of a braille writing tutor to combat illiteracy. In Information and Communication Technologies and Development, 2007. ICTD 2007. International Conference on, pages 1–8. IEEE, 2007.

[13]. A. King, G. Evans, and P. Blenkhorn. Webbie: a web browser for visually impaired people. In Proceedings of the 2nd Cambridge Workshop on Universal Access and Assistive Technology, SpringerVerlag, London, UK, pages 35–44. Citeseer, 2004

[14]. P. Verma, R. Singh, A.K. Singh, V. Yadav, and A. Pandey. An enhanced speech-based internet browsing system for visually challenged. In Computer and Communication Technology (ICCCT), 2010 International Conference on, pages 724–730. IEEE, 2010.