ENGINEERING TERMS TALKING DICTIONARY

Md. Abdullah Al Noman* [1, 2], Rubel Kanti Das [3], A.S.M NOMAN [4]

[1] National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing, China.

[2] School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China.

[3] Department of Electronic & Communication Engineering, Southern University Bangladesh.

[4] Department of Electrical & Electronic Engineering, Southern University Bangladesh.

Article Info

Article history:

Received : April 26, 2023
Revised : May 14, 2023
Accepted : May 20, 2023

Keywords:

Dictionary
Talking dictionary
Raspberry Pi
STT
TTS

ABSTRACT

The present study introduces a practical tool for translating diverse technical terminology in the field of engineering. The device has the capability to comprehend the requirements of its users and, if the relevant information is available within its database, it provides an explanation of the assigned task. Furthermore, in the event that the response requiring a reply is not present within the database, the dictionary prompts the user to indicate whether they wish to document the response. Consequently, the dictionary's database has the potential to be enhanced. The Raspberry Pi has been utilized as the central controller.

The system comprises of two components, namely Speech-to-Text (STT) and Text-to-Speech (TTS).

Corresponding Author:

Md. Abdullah Al Noman
National Engineering Laboratory for Electric Vehicles, Beijing Institute of Technology, Beijing, China
5th South Zhongguancun Street, Haidian District, Beijing, China, Postcode 100081
Email: nomansub14@gmail.com

1. INTRODUCTION

In the field of engineering, it is necessary to acquire knowledge of various technical terminologies. In such instances, relying solely on a dictionary is considered to be a crucial method. The efficacy of utility-based text-oriented dictionaries can be significantly augmented through the utilization of electronic dictionaries that employ speech as the primary mode of interaction. The present study introduces a talking dictionary as a potential substitute for traditional dictionary materials.

This paper presents a talking dictionary that possesses the ability to comprehend user inquiries and respond with a natural-sounding voice [1, 2].

The manuscript is organized into four distinct sections. Section II provides a detailed exposition of the system, while Section III expounds upon the operational methodology. Finally, Section IV summarizes and concludes the investigation.

2. LITURATURE REVIEW

The project utilizes an embedded platform that is based on the Linux operating system. Embedded systems employ the Linux kernel and diverse open-source constituents in Embedded Linux [3]. The operating system in question is constructed utilizing the open source software development model and is primarily employed for server applications. Various categories of Embedded Linux systems exist, such as ETLinux, LEM, LOAF, uClinux, U Linux, and Thin Linux. Several platforms that are commonly used include Google's Android operating system, which is a popular type of embedded Linux found on smartphones, as well as Tizen, an embedded Linux system designed for smartphones. Additionally, Debian is utilized on Raspberry Pi, while Emdebian Grip, BusyBox, OpenMoko, Familiar Linux, and Mobilinux are also commonly used platforms [4].

Several Linux-based embedded devices that are commonly utilized include Arduino Yun, BeagleBoard-xM, BeagleBone A6, Cubieboard, pcDuino, RascalMicro, and Raspberry Pi [5]. In addition to utilizing the Linux operating system, the implementation of speech recognition technology may facilitate ease of use. The present study involves the utilization of Raspberry Pi, which is operated by the Linux operating system. In addition to the utilization of the 'pocketsphinx' library, the STT (Speech to Text) system has been implemented as per reference [6]. Conversely, the installation of python-pymad has been completed for the purpose of implementing a Text-to-Speech (TTS) system [7].

3. OVERVIEW OF THE SYSTEM

The device functions as a speech-enabled lexicon that is capable of engaging in user interaction. In the event that an inquiry is posed and its answer is present in the database, the system will audibly articulate the response utilizing a naturalistic vocal tone and a comprehensible pace. In the event that the queried information is not present in the database, the device prompts the user to indicate whether they wish to document the answer to the question. In the event that the user desires to document their response, the device will proceed to record said response and subsequently store it within the database.

4. WORKING PROCEDURE

The Raspberry Pi has been utilized as the central controller of the device. According to reference [8], the software facilitates the use of multiple programming languages, including but not limited to C, C++, JAVA, and PHP. Upon activation, the device initiates input/output operations and performs an error check. In the event of an error, the I/O is reinitialized. In the absence of any errors, the system commences listening to user input. However, the device is unable to process the inquiry until it receives the pass key. Prior to posing the inquiry, it is necessary for the user to specify the pass key. When the pass key is enunciated, the device initiates STT (Speech to Text) functionality [9]. Subsequently, the apparatus conducts spectral analysis. The database is queried for the STT using spectral analysis. In the event that the query is contained within the database, the system will respond by utilizing a TTS (Text To Speech) mechanism to audibly communicate the answer. The inquiry is conducted through spectral analysis. When the precision of a given spectrum aligns with approximately 80% of the spectra in the database, it is designated as the favored response.

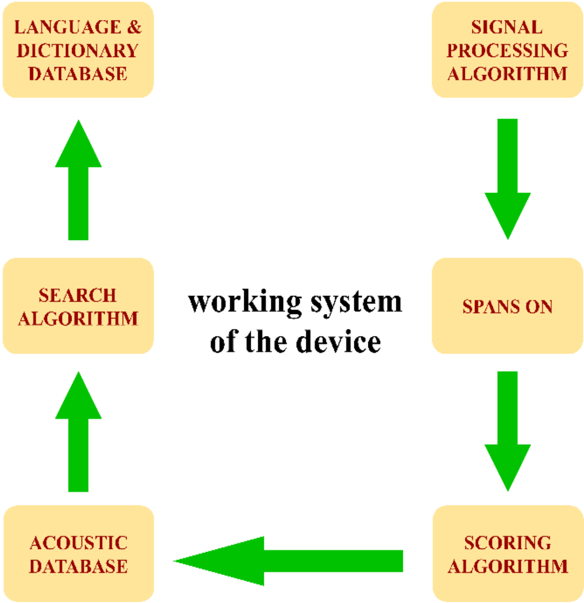

Figure 1: Working system of the device

Figure 1 depicts a succinct overview of the system's functioning. A signal processing algorithm is employed for the purpose of sound processing. Upon completion of sound processing in speech-to-text (STT), the resulting string is assigned an acoustic score, which is utilized for database search purposes. The technique employed in this instance is referred to as the "spans on" method. The process of scoring involves the utilization of a scoring algorithm, which is complemented by an acoustic database to facilitate the identification of corresponding matches. A searching algorithm is employed for the purpose of conducting a search, with the language and dictionary database serving as the underlying database.

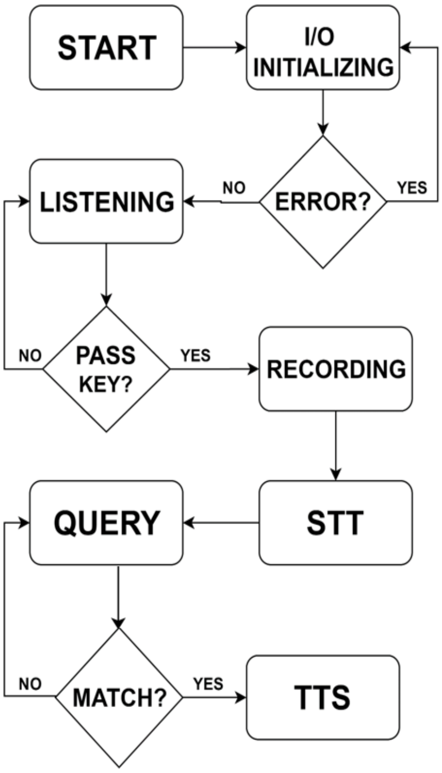

Figure 2: Algorithm for the system

5. APPLICATION DOMAIN

Research has demonstrated that interactive learning is more comprehensible and has the potential to be retained in long-term memory. The utilization of visualization techniques appears to be a more comprehensible approach to the process of acquiring knowledge. The utilization of a talking dictionary can provide an engaging and enjoyable means of learning, particularly in instances where studying from a traditional dictionary may lack appeal. This is due to the interactive nature of the talking dictionary, which fosters a user-friendly experience. The software operates in both offline and online modes.

6. RESULTS AND DISCUSSION

The accuracy of the given task was evaluated through testing the device with three users, as presented in Tables 1, 2, and 3. The rates of acceptance and rejection are also displayed therein. Table 4 presents a comparative analysis of the acceptance rate expressed as a percentage for the spoken language. The obtained acceptance rate is 80% in aggregate.

Table 1: Accuracy testing for spoken terms by trial 1

Trial No:- 01			
Spoken Term	**Frequency**		
	No	**Accepted**	**Not- Accepted**
Electromagnetic	10	9	1
Galvanometer	10	8	2
Earth Ground	10	7	3
Echo	10	10	0
Voltage	10	8	2

Table 2: Accuracy testing for spoken terms by trial 2

Trial No:- 02			
Spoken Term	**Frequency**		
	No	**Accepted**	**Not- Accepted**
Electromagnetic	10	8	2
Galvanometer	10	8	2
Earth Ground	10	6	4
Echo	10	10	0
Voltage	10	8	2

Table 3: Accuracy testing for spoken terms by trial 3

Trial No:- 03			
Spoken Term	**Frequency**		
	No	**Accepted**	**Not- Accepted**
Electromagnetic	10	8	2
Galvanometer	10	8	2
Earth Ground	10	6	4
Echo	10	10	0
Voltage	10	8	2

Table 4: Accuracy (in %) for spoken term

Spoken Term	Trial No: 01	Trial No: 02	Trial No: 03
Electromagnetic	90%	80%	80%
Galvanometer	80%	80%	80%
Earth Ground	70%	60%	60%
Echo	100%	100%	90%
Voltage	80%	80%	80%
Overall Percentage	84%	80%	78%

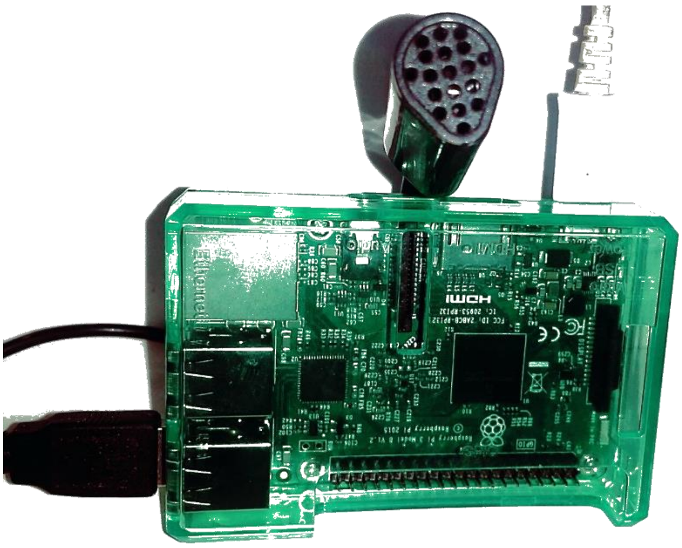

Figure 3: Prototype

7. CONCLUSION

This paper presents a description of a talking dictionary that incorporates interactive features for user engagement. Moreover, it can be easily managed. We are currently engaged in efforts to enhance the level of intelligence by incorporating features such as the ability to comprehend diverse forms of inquiries, as well as the capacity to communicate in various languages.

REFERENCES

[1] S.Suresh, Y. Sindhuja Rao, "Modelling Of Secured Voice Recognition Based Automatic Control System", IJETCSE, ISSN: 0976-1353, Volume 13, Issue: 2 –MARCH 2015

[2] Anusha Bhargava, Karthik V. Nath, Pritish Sachdeva and Monil Samel, "Reading Assistant for the Visually Impaired", IJCET, E-ISSN 2277 – 4106, P-ISSN 2347 – 5161, Vol.5, No:- 2, April 2015

[3] PramodKumar Singh, Prof Swanili Karmore, "Startup Time Optimization Techniques for Embedded Linux" , International Journal of Scientific & Engineering Research, ISSN 2229-5518, Volume 7, Issue 3, March-2016

[4] Embedded system, Wikipedia, USA, 15.01.2017, https://en.wikipedia.org/wiki/Linux_on_embedded_systems

[5] Linux Embedded Devices Comparison: Yun, BeagleBoard, Rascal, Raspi, Cubieboard & pcDuino, cooking hacks, Zaragoza (Spain), 18.01.2017, https://www.cooking-hacks.com/blog/new-linux-embedded-devices-comparison-arduino-beagleboard-rascal-raspberry-pi-cubieboard-and-pcduino/

[6] Prachi Khilari , Prof. Bhope V. P., "Implementation of Speech to Text Conversion" , IJIRSET, ISSN(Online) : 2319-8753, ISSN (Print) : 2347-6710, Vol. 4, Issue 7, July 2015

[7] Suhas R. Mache, Manasi R. Baheti, C. Namrata Mahender, "Review on Text-To-Speech Synthesizer", IJARCCE, ISSN (Online) 2278-1021, ISSN (Print) 2319 5940, Vol. 4, Issue 8, August 2015.

[8] Richard Grimmett, "Raspberry Pi Robotics Project", second edition, 2015, ISBN: 978-1-78528-014-6

[9] Aamir Lakhani, Joseph Muniz, "Penetration Testing with Raspberry Pi", ISBN: 987-1-78439-643-5, January 2015

[10] Peter Membrey, David Hows,"Learn Raspberry Pi with Linux", ISBN: 978-1-4302-4821-7

[11] What to choose linux operating system or windows operating system? , Host Department, USA, 16.01.2017,

[12] https://www.hostdepartment.com/blog/2013/12/27/choose-linux-operating-system-windows-operating-system/

[13] Deepa V.Jose, Alfateh Mustafa, Sharan R, "A Novel Model for Speech to Text Conversion" , IRJES, ISSN (Online) 2319-183X, (Print) 2319-1821, Volume 3, Issue 1 (January 2014), PP. 39-41

BIOGRAPHIES OF AUTHORS

MD. ABDULLAH AL NOMAN was a lecturer in the Electrical & Electronic Engineering Department at Chittagong Institute of Engineering & Technology (CIET), affiliated with Chittagong University. Southern University awarded him a Bachelor of Science in Electrical and Electronic Engineering. He is pursuing his master's degree in mechanical engineering at the Beijing Institute of Technology. His research interests include autonomous vehicles, computer vision, machine learning, and deep learning, among other things. He can be contacted at email: **nomansub14@gmail.com**