Elevator Control Using Speech Recognition for People with Physical Disabilities

Komal Mahajan[1], Riddhi Nahar[2], Dhanali Khairnar[3], Shrutika Kinge[4], Sujata Suryawanshi[5]

[1,2,3,4] Student, Department of Information Technology, MVPS's K.B.T. College of Engineering, Nashik, INDIA

[5] Professor, Department of Information Technology, MVPS's K.B.T. College of Engineering, Nashik, INDIA

Abstract

This proposed system that will be useful for people with incapacities, to be specific lifts constrained by voice. Disability is someone who does not have a complete hand organ or hand organs but does not function properly, but the person can still use voice to control the elevator. The research combines speech recognition technology with machine learning to make the elevator control equipment that can be controlled by voice. Speech Recognition is a system that functions to convert spoken language into input data. The system input is human speech. The system will identify spoken words to input data for control equipment. Control of this hardware requires an exact word and can only recognize some words. These proposed systems are usually more accurate and more easily trained but could not recognize words that are beyond vocabulary ever taught. This system uses a python programming language and machine learning libraries with Raspberry pi 3B+. The word can be trained for more than hundreds of words.

Keywords: Raspberry Pi 3B+, speech recognition, elevator, Machine Learning.

I. INTRODUCTION

In this rapid world of innovation where a voice starts, its time of domination to replace the touch screens from smartphones to huge computer systems, bringing voice in day-to-day affairs becomes significant. Elevators being one such system used in daily life serves this purpose of making future generations hands free which also becomes a boon for the disabled.

The essential working guideline of the lift depends on the elevator algorithm, where a lift can choose to stop dependent on two conditions. The first one being the direction and the second one based on the current floor and destination floor. The elevator is generally made up of cables, motor, pulleys based on traction, climbing, or hydraulic system. To serve project purposes it can also be designed by connecting the elevator system to a raspberry pi to accept input voice. The voice control option is attractive for several reasons. It is potentially appropriate for many elevator users since the system can be used by any individual capable of consistent and distinguishable vocalization [7]. Voice control also reduces physical requirements. However, the recognition accuracy of the Automatic Speech Recognition (ASR) system is a constraint in the deployment of many voice-controlled systems in a real-world application.

In this proposed architecture, a voice-operated elevator system is proposed where the user's input commands to control the movement of the elevator system are kept convenient for the users. The commands include voice input for the floor operations, directions, elevator car's door operation, and a special option to place a call of speaker's choice in case of any unexpected event that requires immediate action.

The paper is organized as follows: Section II mentions briefly a review of some of the related earlier works. Section III describes proposed architecture for this project and explains various features of speech recognition. Section IV explains the experimental setup for the proposed model. Section V discusses the experimental results that are obtained on performing the tests under laboratory conditions.

II. LITERATURE REVIEW

Speech recognition had been effectively contemplated since the 1950s, however late improvements in Personal Computer (PC) and telecommunication technology have enhanced speech recognition abilities [1]. From a valuable perspective, discourse acknowledgment has handled issues, improved advantages, and purchased a more prominent transformation in the current situation [1]. Voice control can replace the function of a push-button efficiently [2]. Speech recognition is a very complex issue. It includes numerous calculations that oblige high computational necessities [2]. The mixture of the utilization of programmed speech recognition systems, human PC interfaces, telephony, or robotics has driven the exploration of a vast academic group over decades ago [3]. Automatic speech recognition is used in various areas ranging from medical transcription to game control, from call center dialogue systems to data recovery [4].

This automatic speech recognition process is used in most voice-controlled systems. The voice-controlled wheelchair for the physically challenged was proposed in 2002 [7]. This paper had portrayed an investigation that thought about the presentation of abled and impaired individuals utilizing voice control to work a force wheelchair both with and without route help, where the navigations were assisted by the sensors to identify and avoid obstacles in the wheelchair's path [7].

The intelligent lift control model that uses voice recognition was proposed in 2003 [8]. This proposed elevator system had been controlled by speech and sensor panel [8]. The modification of the well-known DTW (Dynamic Time Warping) algorithm was used [8]. The set of various voice commands for the model consisted of eight Lithuanian words [8]. The model was specifically designed for domestic use (smart houses).

III. PROPOSED ARCHITECTURE

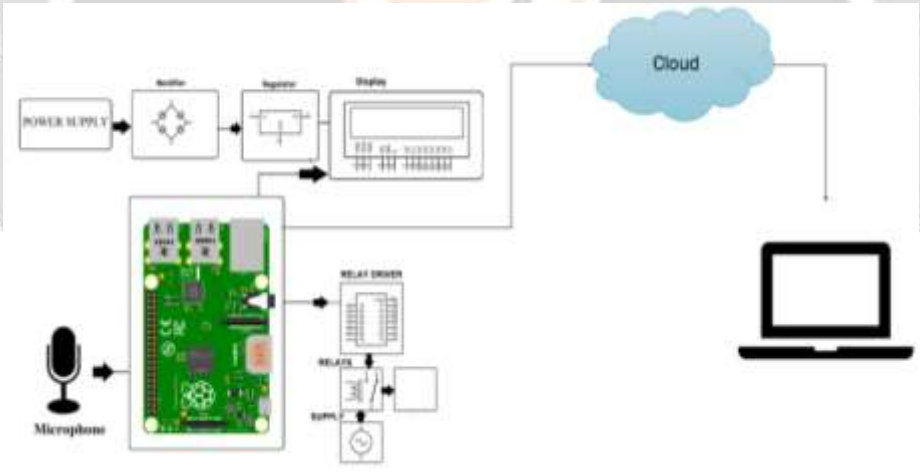

Figure 1 System Architecture

In this proposed architecture, a voice-controlled elevator system is introduced where the input commands to simulate the movement of the elevator system are kept convenient for the users. The commands include voice input for the floor operations, directions, elevator car's door operation, and a special option to place a call of speaker's choice in case of any unexpected event that requires immediate action. Raspberry pi 3B+ with Linux operating used as main controlling unit and USB microphone is used get user input.

The input voice to the Automatic Speech Recognition (ASR) system deduces a command for the control device. This device controls entire elevator operations including its speed and traction. It also handles any emergency. In the event of unexpected situations, to avoid panicking, calling methods have been improvised where the user can verify the number using a TTS system. Any cellular number of speaker's choice can be called, Otherwise, the normal lift operations are resumed.

Sphinx4 is a Java speech recognition library where Speech recognition used is as an Open-Source recognition platform from CMU. It gives a quick API to change the speech recordings into text with the help of "CMUSphinx" acoustic models. It can be utilized on servers and as a part of desktop applications and it is highly portable. Besides speech recognition, Sphinx4 serves to distinguish speakers, adapt models, allow highly flexible user interfacing, and more. Sphinx4 supports US English and many other dialects.

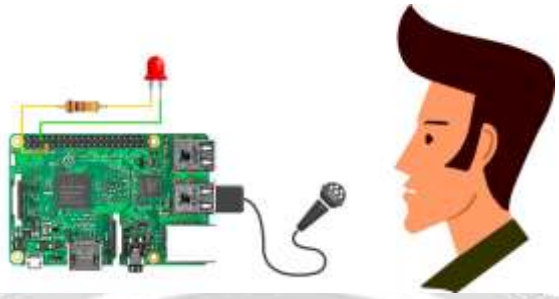

Figure 2 Raspberry pi Speech recognition

The recognition platform of Sphinx4 helps to add library files into dependencies of the project and there are a few high-level recognition interfaces such as Live Speech Recognizer, Stream Speech Recognizer, Speech Aligner, and Speech Result. These interfaces along with the Acoustic model, Dictionary, Grammar/Language model, and Source of speech are executed for the task of speech recognition.

The four necessary attributes of Sphinx4 are described in detail. The acoustic model is used as part of automatic speech recognition to represent the association between an audio signal and the phonemes. Some context-independent models contain properties (feature vectors for each phone) and context-dependent ones (created from phones with context) [10]. A phonetic dictionary has a mapping from words to phones, which is not very effective. For instance, just two to three pronunciation variants are noted in it, yet it is sufficient most of the time. The dictionary is not the only variant of mapper from words to phones. It might be possible with some complex functions such as machine learning calculation [10].

A language model is used to limit word search and to represent measurable properties of speech [10]. It characterizes which word could take after already perceived words and serves to limit the coordinating process by stripping words that are not likely. To achieve a better recognition rate, a language model must be effective in search space restriction [10]. This implies that it should be very good at anticipating the next word. A language model normally allows the vocabulary to contain the words specific to that dialect. That is an issue for name recognition. To manage this, a language model can contain smaller chunks like sub-words or even phones. The search space restriction for this situation is generally worse and corresponding recognition accuracies are lower when compared with a word-based language model. Source of speech in Sphinx4 generally is open source where any speaker can pronounce any word that has been already defined in the phonetic dictionary belonging to any one of the specified languages.

Text-To-Speech (TTS) synthesizer is a computer-based system that reads any text aloud, whether it was introduced in the PC by an administrator or checked and submitted to an Optical Character Recognition (OCR) system [11]. It is a speech synthesis application that is used to make a spoken sound version of the text in a computer report, such as a help document or a Web page [12]. It can empower the reading of computer display information for the visually challenged person or may just be utilized to enlarge the perusing of an instant message.

	Commands	*Description*
Floor numbers	*One, two, three, four, five, six, seven, eight.*	*The simulated elevator moves to the corresponding floor number.*
Direction	*Up, down.*	*The simulated elevator moves along the corresponding directions.*

Door operations	*Open, close*	*The opening and closing of the elevator car's door are controlled.*
Cellular number	*The sequence of ten digits**	*Call for help is placed to this number*

Table 1 - elevator database

IV. CONCLUSION

This paper explains how voice control can become a boon in the future in everyday life utilizing an elevator simulation. The proposed work shows the feasibility of developing an elevator system based on voice control. It also incorporates a voice feedback system in case of placing emergency calls, which helps the user to verify the correctness of the number. In the future, the size of the experiment can be improvised to make this model a real-time system. Also, the recognition accuracy can be improved by making the speech recognition system speaker-dependent and by including the aspect of robustness to noise.

V. REFERENCES

[1] K. Michael, M. K. Anna, T. Justin, andY. Tony, "Development of a Voice Recognition Program", May 2002.

[2] P. Martin, P. Dušan, and F. Peter, "Speech control for car using the TMS320C6701 DSP", 10th International Scientific Conference, Bratislava, pp. 97-100, 2000.

[3] A. V. Nefian, L. Luhong, P. Xiaobo, X. Liu, C. Mao, and K. Murphy, "A coupled HMM for audio-visual speech recognition", IEEE Proceedings (ICASSP '02) on Acoustics, Speech, and Signal Processing, Volume 2, pp. 2013-2016, 2002.

[4] H. Thomas, E. H. Asmaa, N. W. Stuart, and. Vincent, "Automatic speech recognition for scientific purposes – web as".

[5] T. Q. Muhammad and A. A. Syed, "Voice-controlled wheelchair Using DSK TMS320C6711", 2009 International conference on signal acquisition and processing, pp. 217-220.

[6] 2012 International Conference on Computer Science and Electronics Engineering, Youhao YuDepartment of Electronics and Information Engineering, Putian University, Putian, Fujian, 351100, China "Research on Speech Recognition Technology and Its Application".

[7] C. S. Richard and P. L. Simon, "Voice Control of a Powered Wheelchair", IEEE transactions on neural systems and rehabilitation engineering, vol 10, no 2, pp. 122-125, June 2002.

[8] P.Cernys, V.Kubilius, V.Macerauskas, and K.Ratkevicius, "Intelligent Control of the Lift Model ", IEEE intentional Workshop on Intelligent Data Acquisition and Advanced Computing System: Technology and Applications

[9] Sphinx4 dependencies URL:http://cmusphinx.sourceforge.net/wiki/tutorialsphinx 4

[10] Sphinx4 attributes URL: http://cmusphinx.sourceforge.net/wiki/tutorialconce pts

[11] Text-to-Speech(TTS) URL:http://searchmobilecomputing.techtarget.com/definit ion/text-to-speech