International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 4144 – 4151

Implementation of Speech Recognition Based Eye Vision Test

Anurag Potharaju[1],Amit Ghogre[2],Harsh Kurulkar[3],P.G.Chilveri[4]
Department of E&TC Engineering, SKNCOE, SPPU, Pune, India [1]anuragpotharaju@gmail.com
[2]amitghogre7@gmail.com [3]harshkurulkar7@gmail.com [4]pgchilveri@gmail.com

***Abstract*: -**
In today's world, many underdeveloped countries often neglect eye care due to lack of awareness. Our idea describes a speech recognition-based eye test which is capable of performing the entire Snellen eye test on its own and estimates one's visual acuity. The test uses a few random letters like "A", "E", "I", "R", and "L" with the same 6 possible text sizes for each of these letters. The user speaks his/her guess into the microphone. The speech recognition portion uses energy thresholding. The system will then determine the next step based on whether the user guessed the displayed letter correctly or not. If the user guesses enough correctly, the text size will continue getting smaller and If the user guesses too many letters incorrectly, the system will display the result corresponding to the current text size. The thresholds for amount guessed and amount guessed correctly are set in the code. The software used for this Project is Python for Raspberry Pi

Keywords: *Speech Recognition,Visual acuity, Snellen Chart,Raspberry Pi*

I. INTRODUCTION

The standard Snellen letter chart based diagnostic system does not work always. There are instances where the first few students who undergo this diagnostic system, memorize the letter sequence of the Snellen letter chart and convey the same to other fellow students. Hence other students simply read out the Snellen letter sequence from their memory, not by looking at the Snellen letter chart. Thus, there is a need of randomizing the sequence of letters being displayed on the Snellen letter chart for every student to be diagnosed for eye testing. Our idea describes a speech recognition-based eye test which is capable of performing the entire Snellen eye test on its own and estimates one's visual acuity. The test uses a few random letters like "A", "E", "I", "R", and "L" with the same 6 possible text sizes for each of these letters. The user speaks his/her guess into the microphone. The speech recognition portion uses energy thresholding. The system will then determine the next step based on whether the user guessed the displayed letter correctly or not. If the user guesses enough correctly, the text size will continue getting smaller and If the user guesses too many letters incorrectly, the system will display the result corresponding to the current text size. The thresholds for amount guessed and amount guessed correctly are set in the code

Eye is an important organ of human body and regular eye tests are important because our eyes don't usually hurt when something is wrong. Now a days, for a simple eye check-up a person needs to take an appointment beforehand, sometimes it takes days to make an appointment. Such delays can be mitigated if a smart phone application or computer-based program is available which can estimate eye vision power. The condition in the developing countries is even worse. In the developing country, the doctor-to-patient ratio is quite low. Moreover, people having low income show negligence for a regular eye check-up. But regular eye check-up is must as without this many problems may occur such as a driver with lower eye sight may cause accidents. Low myopia may turn to severe myopia which has higher prevalence of coexisting disease and complications. For this

ISSN: 2233-7853 IJFGCN

reason, a low-cost automated eye check-up system is needed.

II. EVIEW OF LITERATURE

The literature survey of the vision based systems are reviewed in this section. Shikha Gupta et al.[1] proposed paper finding different speech aware applications available in the market. In this paper, the authors look for various efficient alternative methods like speech recognitions systems (SRS) for devices where typing becomes difficult. Author has used embedded platforms to make the SRS. They use speech analysis technique for the SRS where analysis, extraction of the information of the speaker identity is done .In this paper, various techniques are discussed about SRS. This paper also presents the list of technique with their properties of Feature extraction and Feature matching. The goal of speech recognition is to analyses, extract, characterize and recognize information about speaker identity. Variety of the techniques are used for determining the speech characteristics. Also, author specified different types of Speech Analysis techniques which are i) Segmentation analysis, ii)Sub segmental analysis, iii)Supra segmental analysis Lalima Singh et al[2] proposed a paper on "Speech Signal Analysis using FFT and LPC", in this paper she has done speech analysis using FFT and LPC . She proposed that in speech processing there are different signal analysis techniques are used. In this paper two most commonly used signal analysis technique are used. First is Fast Fourier Transform(FFT) and second is Linear predictive Coding(LPC). These techniques are used to extract and compress some features of speech signal for further processing. In this paper five samples of single word is taken by same person. These samples are analyzed using FFT and LPC in MATLAB and spectra plus software. After analysis different parameters of samples are obtained for FFT and LPC spectrum individual, also this paper analyses the accuracy of feature extraction based on modelling which is implemented using Mel Frequency Cepstral Coefficient (MFCC) and Hidden Markov Model (HMM) for two different type connected and continuous speech. In this paper two most commonly used signal analysis technique are used. First is Fast Fourier Transform (FFT) and second is Linear predictive Coding (LPC) which are used to extract and compress some features of speech signal for further processing. In this paper, we have used two signal analysis techniques FFT and LPC. Using these techniques, the speech signal first analysed and different parameter of speech. Neha Jain et al.[3] proposed a research on Speech Recognition Systems. In This Paper, she has done has the detailed study of the mechanism, the challenges and the tools to overcome those challenges with a concluding note that would ensure that with the advancements of the technology which this paper has reviewed various aspects related to Speech Recognition Technology and the systems implementing this technology to engineer Speech Recognition Systems. The paper first describes what actually lead to the emergence of developing Speech Recognition Systems, continuing with how the mechanism of conversion of speech takes place in distributed real-time systems. Finally, this paper reviews the advances that have taken place after the development of traditional Speech Recognition Systems, also this paper briefly makes a quick comparison between the algorithms and the models that were and now are being used to implement these Speech Recognition Systems. Lindasalwa Muda et al[4] presented a mobile healthcare (mHealth) system for estimation of visual impairment that provides easiness by specifying the degree of an eye as orthoscopes. Our proposed system called AcuMob which is an Android based mobile application aimed to be used by patients who have myopia. In this, digital signal processes such as Feature Extraction and Feature Matching are introduced and studied to represent the voice signal. Several methods such as Liner Predictive Coding (LPC), Hidden Markov Model (HMM), Artificial Neural Network (ANN) and etc. are evaluated with a view to identify a straight forward and effective method for voice signal. The results show that these techniques could be used effectively for voice recognition purposes. Several other techniques such as Liner Predictive Coding (LPC), Hidden Markov Model (HMM), Artificial Neural Network (ANN) are currently being investigated. Akhan Muhammed Ali et al.[5] proposed paper on

ISSN: 2233-7853 IJFGCN

mobile healthcare (mHealth) system for estimation of visual impairment that provides easiness by specifying the degree of an eye as orthoscopes. The system is called AcuMob which is an Android based mobile application aimed to be used by patients who have myopia. The system is developed with using Xamarin framework and voice commands are used to interact with mobile app. Some preferable letters that are suggested by the ophthalmologists were used in the system .This research has aimed at making a prediction about the visual acuity of users via a mobile application. As part of future work, the necessary arrangements will be made so that AcuMob can also run on the iOS platform. Lütfü SARIBULUT et al.[6] proposed paper on FFT transform . The objective of this paper is to study about FFT and learn to classify and analyses various types of PQ disturbances in power systems and signal process. In this paper, the fundamentals of Fourier series and FT with their types and distinctions over other disturbance detection methods are clearly discussed. This paper presents the fundamentals of Fourier series, Fourier transform, discrete Fourier transform and fast Fourier transform with simple examples and review of Fourier transform to provide a clear understanding of its applications in power quality issues. Aakash Agarwal et al.[7] in which they developed an eye test on android .The objective of this paper is about the android app which is called Dr. Eye which is an android application that focuses on the calculation of the vision acuity of a patient in a similar way an ophthalmologist. The use of front camera and the speech to text conversion of the Android API makes it possible to achieve this process. This paper also predicts what defect a person can have based on the answers of a questionnaire as well as provides the corrective measures and treatments for the defect and also displays the nearest Ophthalmologist or the nearest Hospitals to the user. In future, this application can be extended by improving existing database for more defects of the eye and integrating proper reasoner for predicting defects and corrective measures. Manjutha M et al.[8] proposed a paper where they have developed Automated Speech Recognition System .This paper analysis the accuracy of feature extraction based on modelling which is implemented using Mel Frequency Cepstral Coefficients (MFCC) and Hidden Markov Model (HMM) for two different type connected and continuous speech. In this paper, continuous and connected words are considered and MFCC features were extracted from the speech corpus. The extracted speech signal is trained by HMM model. In this paper, MFCC feature is extracted and the speech is trained by HMM model which is implemented for both connected and continuous speech. In order to improve the accuracy, other modelling techniques will be implemented in future. Saud Khan et al.[9] proposed a technique that aims at the ease of accurately testing vision anywhere. In this method, author has made use of two speech extracting techniques which are FFT and MFCC and also given information about the same. Snellen chart and Ishihara tests were performed using screens and a microphone for the patient to utter their observations from the screen. After the proposed feature extraction and classification techniques, the model predicts the correct or wrong guesses of the user sequentially and then displays the eye test results on screen. Annisa Novantina et al.[10] proposed a paper on which Snellen Chart based eye test is created using android control. This study has shown the development of the android Snellen control chart to facilitate the visual inspection process of the eye. In this study, an examination and testing of the Android Snellen chart control were carried out. Tests were conducted to determine the ability of Bluetooth HC-05 to send data to android at a distance of 10 meters. The test is carried out with 2 conditions, namely the first condition in one room without a barrier while the second condition is different from the room with a change in range every 2 meters. Tests for each distance range are carried out 5 times. This research was made with the Atmega 16 microcontroller using the 2803 ULN and Bluetooth HC05 lamp drivers to connect with Android. R N Adnan et al.[11] proposed this paper where their objective is to use Snellen chart to measure persons visual acuity consisting series of black capitalized letters on a whiteboard arranged in row. The control device is an application on a cellphone connected via Bluetooth by utilizing LED as lighting while showing

ISSN: 2233-7853 IJFGCN

letters on the Snellen chart. It will facilitate ophthalmologists to perform examinations. In this paper It can be concluded that the electronic Snellen chart is able to be created by the control through the application on a smartphone connected to Bluetooth. The electronic Snellen chart can function adequately, it is appropriate and suitable for a checkup. Tingxiao Yang[12] worked on investigating the algorithms of speech recognition. The author programmed and simulated the designed systems for algorithms of speech recognition in MATLAB. There are two systems designed in this thesis. One is based on the shape information of the cross-correlation plotting. The other one is to use the Wiener Filter to realize the speech recognition. The simulations of the programmed systems in MATLAB are accomplished by using the microphone to record the speaking words. The simulations of the programmed systems in MATLAB are accomplished by using the microphone to record the speaking words. SANJA GRUBESA et al.[13] proposed this paper where they discussed about method of speaker recognition based on wavelet functions and neural networks is presented in this paper, the method is based on spectral analysis of time windowed speaker's voices, using the Fast Fourier Transform (FFT) and creating the averaged spectrum over the defined time. Pranab Das et al[14] proposed paper in which he described and performed the Voice Recognition System. In this paper the fundamentals are discussed and its recent progress is investigated. The various approaches available for developing a Voice Recognition System based on adapted feature extraction technique and the speech recognition approach for the particular language are compared in this paper. The main aim of our project is to develop a system that will allow the computer to translate voice request and dictation into text using MFCC and VQ techniques. Feature extraction and feature matching will be done using Mel Frequency Cepstral Coefficients and Vector Quantization technique. The extracted feature will be stored in .mat file. A distortion measure which is based on minimizing the Euclidean distance will be used while matching the unknown speech signal with the database of the speech signal. In near future, home automation will be completely based on Voice Recognition System

III. METHODOLOGY

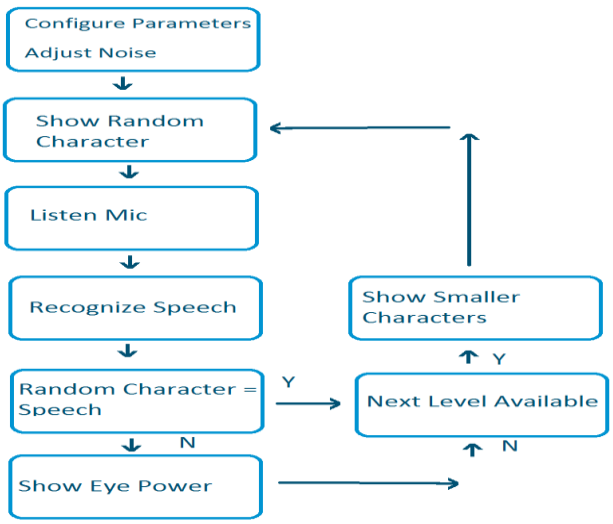

Fig.1BlockDiagramof Speech Recognition Eye Vision Test

First, we have adjusted the parameters which meets the requirements of the mic which you are using like Chunk_size and sample_rate. Chunk_Size - The microphone audio is record in chunks of samples.Higher chunk_size values help avoid triggering on rapidly changing ambient noise, but also makes detection less sensitive. This value, generally, should be left at its default.Sample_rate -

ISSN: 2233-7853 IJFGCN

Higher sample_rate values result in better audio quality, but also more bandwidth (and therefore, slower recognition). Additionally, some machines, such as some Raspberry Pi models, can't keep up if this value is too high. For generating random characters,after setting the parameters, we have to set the characters which we are guessing in the test these characters are from any letter in between A to Z. To listen for input through Mic. After setting random characters, the next process is to listen to the user's words (User has to stand at least 20ft away from the screen). If character shown on screen is A then user has to identify the character and then speak in the mic.Next step is to recognize the speech which user has said and check whether it matches with the generated character. This process is done by Recognizer (). Creates a new Recognizer instance, which represents a collection of speech recognition settings and functionality.Recognizer instance.energy_threshold = 300 represents the energy level threshold for sounds. Values below this threshold are considered silence, and values above this threshold are considered speech. This is how system recognizes speech. If the random character generated on the display is recognized correctly then the size of the next character generated will reduce and the process repeats itself until either the user guesses the character incorrectly or guesses the character correctly in all sizes.If the user guesses the character incorrectly then system will stop the process and will check that particular size of the character that the user has guessed wrong and generate a power in that range.

IV. EXPERIMENTATIONANDRESULTS

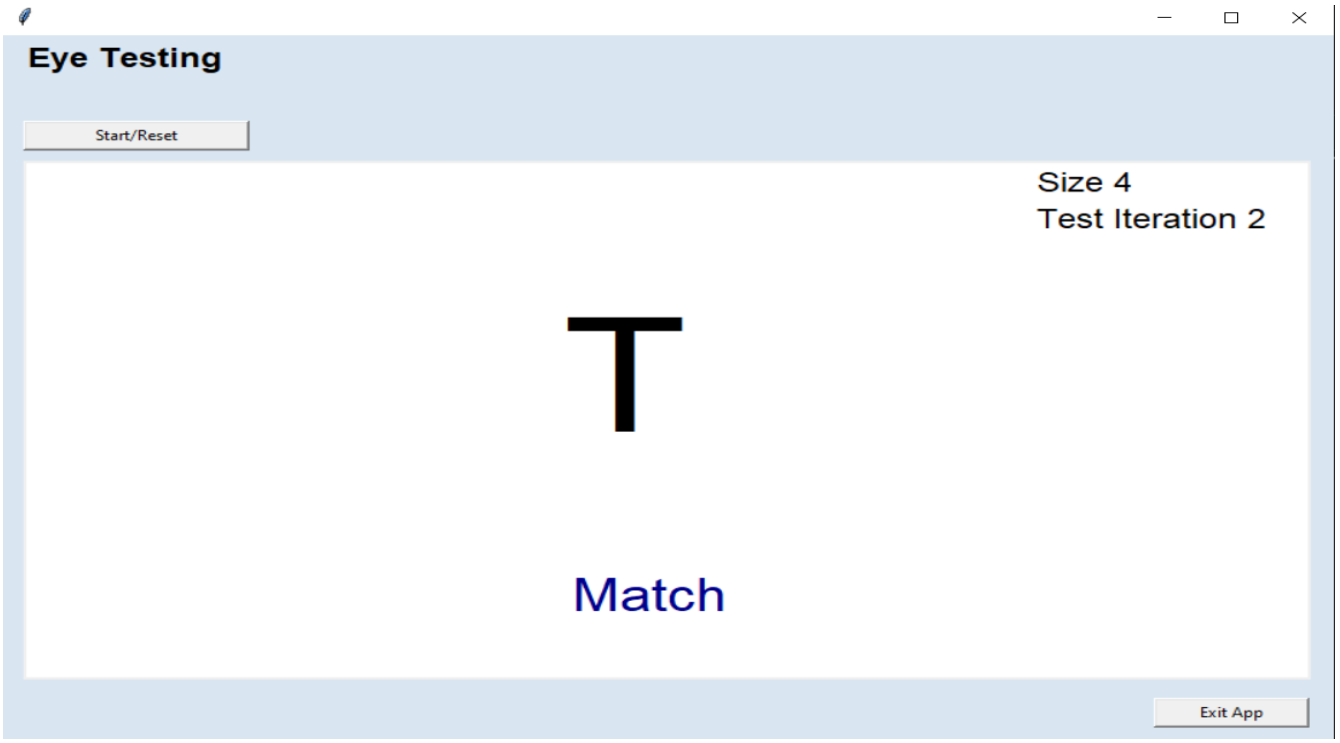

Fig 2 – Output if user's voice and generated character are same

ISSN: 2233-7853 IJFGCN

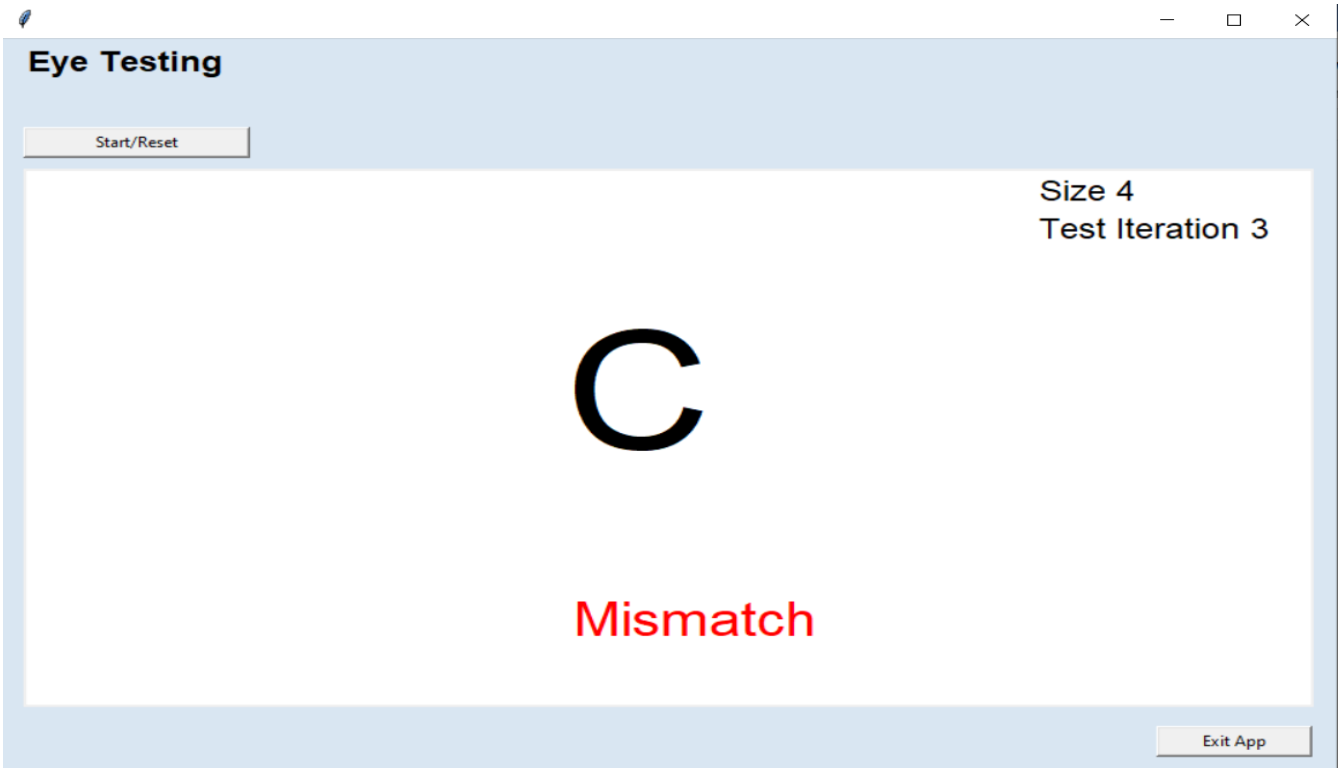

Fig 3 – Output if user's voice and generated character are different

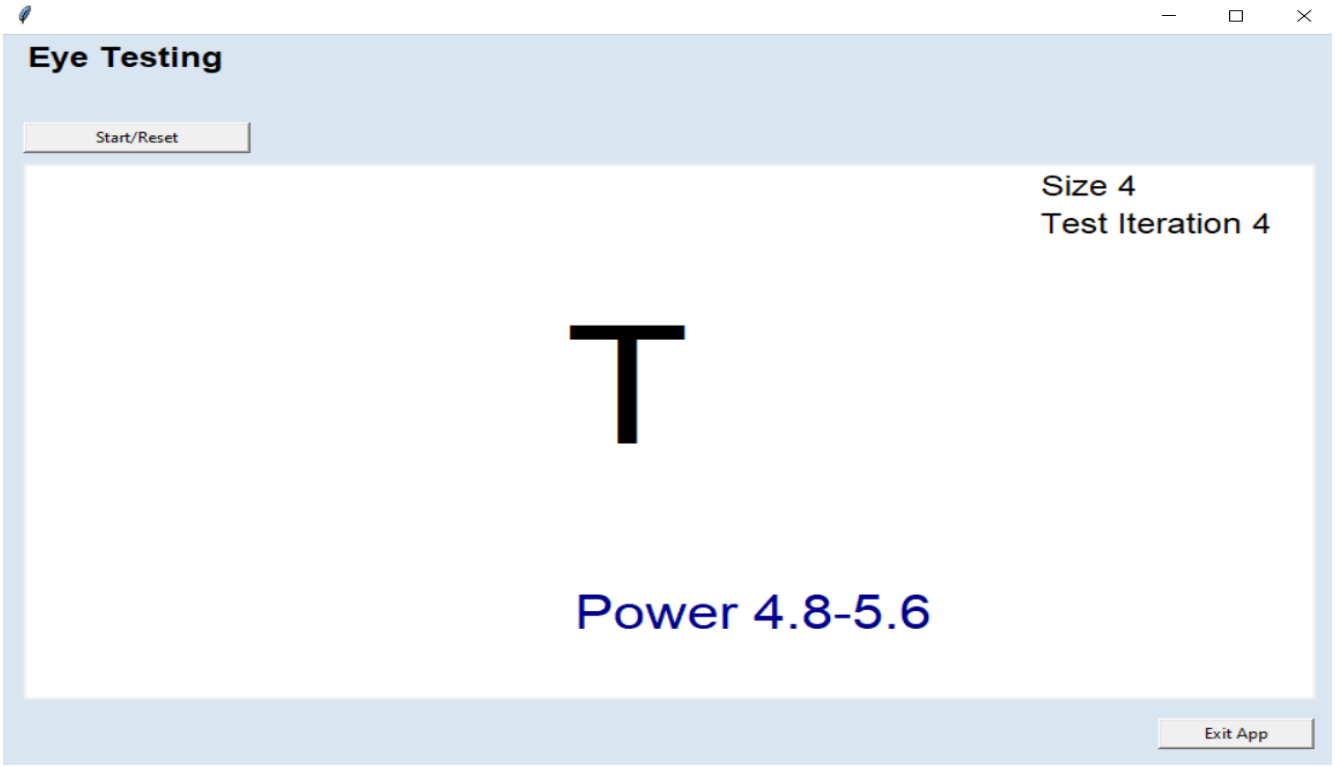

Fig 4 – Final Output indicating user's eye power if user guesses the wrong character more than twice at a specific size level.

To summarize above 3 figures, in our UI we have programmed in such a way that there are 11 sizes

ISSN: 2233-7853 IJFGCN

and at each size there are 5 iterations. A user will go from size 1 to size 2 if he has guessed the characters correctly at least 4 out of 5 times at size 1. If the user guesses incorrectly twice at size 1 then the power corresponding to size 1 will be displayed on the screen. The maximum size is 11 and users guesses correctly at size 11 then the power displayed will be 0.

V. CONCLUSION

This work has implemented Eye Vision Test Using Speech Recognition System in Python. Python speech recognition libraries had provided maximum range of visual acuity with statistical approach in speech recognition model training provided with numerous data. The results show reasonably good success in recognizing continuous speech from various speakers, for a large vocabulary.

VI. SCOPE FOR FUTURE ENHANCEMENTS

Visionary problems such as glaucoma, black eye and similar tests can be added and be successfully identified and, in some cases, may be treated as well before permanent damage. The proposed method can be installed in smart homes as part of Internet of Things where the results can be sent directly to the personal practitioner for further analysis. It can also be installed in different Basic Health Units in remote areas where the need of waiting for a practitioner will be drastically reduced. The computation power is low to the point that it can be run on a Raspberry Pi which further helps in reducing its overall cost.Due to current circumstances and unavailability of hardware and lab, we were unable to perform this project on hardware but it is highly possible to perform this project on Raspberry Pi.

REFERENCES

1. Shikha Gupta, Amit Pathak, Achal Saraf, “A STUDY ON SPEECH RECOGNITION SYSTEM”, 2014
2. Lalima Singh, “Speech Signal Analysis using FFT and LPC”, 2015
3. Neha Jain, Somya Rastogi, “SPEECH RECOGNITION SYSTEMS - A COMPREHENSIVE STUDY OF CONCEPTS AND MECHANISM”, 2019
4. Lindasalwa Muda, Mumtaj Begam, I. Elamvazuthi, “Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques”, 2010
5. Akhan Muhammed Ali, Abdül Halim, “Mobile visual acuity assessment application: AcuMob”, 2017
6. Lütfü SARIBULUT, Ahmet TEKE, Mehmet TÜMAY, “Fundamentals and literature review of Fourier transform in power quality issues” 2013
7. Aakash Agarwal, Kumar Abhishek, Vivek Kumar, Vineet Kumar, Nikhil Prasad, M. P. Singh, “Dr. Eye: An Android Application to Calculate the Vision Acuity”, 2015
8. Manjutha M, Gracy J, Dr P Subashini , Dr M Krishnaveni, “Automated Speech Recognition System”, 2017

ISSN: 2233-7853 IJFGCN

9. Saud Khan, Khalil Ullah, “Visual Acuity Test for Isolated Words using Speech Recognition”, 2014
10. Annisa Novantina Eka Fitria, Priyambada Cahya, Nugraha, Lamidi, “Snellen Chart Based on Android Control” 2015
11. R N Adnan, M Irfan, I P Sari, “Electronic Snellen Chart with Bluetooth Connection and Smartphone App”, 2018
12. Tingxiao Yang , “The Algorithms of Speech Recognition, Programming and Simulating in MATLAB”, 2012
13. SANJA GRUBESA, TOMISLAV GRUBESA, HRVOJE DOMITROVIC, “SPEAKER RECOGNITION METHOD COMBINING FFT, WAVELET FUNCTIONS AND NEURAL NETWORKS” 2008
14. Pranab Das, Vijay Prasad “Voice Recognition System: speech-to-text”, 2005

ISSN: 2233-7853 IJFGCN