**Testing of Three Visual Parameters: Visual Acuity, Contrast Sensitivity and Color Vision with the Implementation of Speech Recognition for Visual Acuity and Contrast Sensitivity Testing**

by

**Trina Ritchell R. Aquino/ECE-4**

**Kenneth D. Chan/ECE-4**

**Heinz Alfred Q. Juan Tong/ECE-4**

Bachelor of Science in Electronics Engineering, 2018

A Thesis Proposal Submitted to the School of Electrical Electronics and

Computer Engineering

in Partial Fulfillment of the Requirements for the Degree

Bachelor of Science in

Electronics Engineering Program

Mapúa University

January 2019

**Chapter I**

# **INTRODUCTION**

There are five major senses that humans possess: sight, smell, taste, touch and hearing. Being organisms who thrive with the usage of the different information gained from their surroundings, humans most definitely need to keep such senses in their best condition possible. According to D.C. van Essen’s study in 2004, the eyes that grant the sense of sight is arguably the most important organ in the human system. The eyes, however, are prone to disorders that are either inborn or are acquired over the span of one’s life. So much that according to the World Health Organization [WHO], there are approximately 1.3 billion people on a world scale that have some form of vision impairment; but thankfully, about 80% of them can be avoided. The National Eye Institute [NEI] stated that, for diabetic retinopathy, early detection reduces risk of blindness by 95%. It is thus highly important to detect as soon as possible and/or monitor eye disorders so that it may be addressed immediately by specialists. There are different eye disorders such as blurry vision, color blindness, poor contrast vision etc. that can hinder the everyday lives of people. Furthermore, some of these disorders such as blurry vision and bad contrast vision can be telling signs of other diseases such as cataracts, glaucoma and diabetic eye diseases. Each disorder can be identified using different tests that utilizes specific equipment and a specialist to operate it; only a few can test multiple disorders using one system. This means that for every patient that needs a check-up, at least one specialist will be occupied for a set amount of time. This amount of time varies depending on the type of test; the procedure; and the equipment used. Most of the equipment used are also bought from outside of the country. With there being less ophthalmologists and optometrists in comparison to the people who have possible eye problems, more specifically so in rural areas compared to the metropolitan of the Philippines, there is a need for locally made systems that can conduct multiple visual parameter preliminary tests for patients, with or without the specialists, and the doctors can perform the in-depth analysis on the patients whose results indicate possible eye diseases later. “Locally made”, as defined by the researchers, means that the system can be built using the prescribed materials by the study, or suitable alternatives, rather than importing the entire working system altogether from outside of the country.

The study by D. C. van Essen in 2004 showed that for humans, 27% of the cerebral cortex is predominantly visual in function, whereas auditory and somatosensory are only 8% and 7% respectively [1]. 80% of the 1.3 billion people in the world having some form of visual impairment or disease can be avoided according to the WHO, 2018 [2], this is because these diseases such as glaucoma, cataracts and other diabetic eye diseases can be detected early through the standard vision screening like visual acuity, contrast sensitivity [3] and even for color vision [4]. NEI in 2015 stated that for diabetic retinopathy, early detection and treatment can cut the risk of blindness by 95% [5]. Thus, there have been numerous studies that propose effective testing procedures for different visual impairments [6]–[9], or low-cost systems that do not require doctors or specialists to operate in order to perform multiple visual screening tests [10], [11]. These are especially useful for those who live in rural areas where there are only a few eye doctors are present. In the Philippines, it is recorded that only 30% of the registered ophthalmologists are scattered around the country while the rest are registered to work in Metro Manila [12]. A study by Huang, Y., Ropelato, S. and Menozzi, M. in 2015 proposed a Vision-screening system that is both fully automatic and computerized. The screening system included tests for visual acuity, color vision, stereo vision, contrast sensitivity as well as peripheral vision. All tests apart from the one for peripheral vision were done using Landolt-C optotypes in order to save time by only mentioning one type of instruction for all four tests: that is to “report the orientation of the opening of a Landolt ring with the input device”. Apart from being effective, the system also managed to do efficient tests at an average rate of 8 tests per 15 minutes [10]. Another study by Paglinawan et al. in 2017 proposed an accurate and low-cost way of screening and detecting strabismus, visual acuity and blind spots using a Raspberry-Pi (R-pi) in order to aid medical authorities. The proposed system utilized the Hirschberg test to detect strabismus; used Landolt-C for visual acuity; and for the visual field test for blurry vision. With the test results being well within the acceptable range of errors, it was concluded that it is viable to create a device for detecting different visual impairments [11]. Another study by Lazaro et al. in 2018 utilized speech recognition through Hidden-Markov model for controlling the optotype characters of the Snellen chart as well as the patient’s input for the visual acuity test [13].

One of the proposed methods [11], was made to identify and diagnose for strabismus and did not include test for contrast vision. The contrast vision as stated prior, is one of the important parameters to detect early signs of eye diseases that may cause temporary and permanent blindness [3] and thus it is important to include such parameter. Although another system did have contrast sensitivity test and four other tests in the standard vision screening tests included, the system ultimately cost four thousand USD due to the laptop and printer that came with it [10]. It is undoubtedly cheaper than the usual system used by hospitals, and possessed a high screening speed, but the price of the system made it hard for users other than large firms and schools, such as the rural areas and the less fortunate population, to acquire. Furthermore, with the assumption that the patient does not have good vision, it is more practical if the test system does not rely too much on vision for the instructions as well as the prompt for input and the input mechanism itself. This was explored by another study for visual acuity testing [13], however, they have recommended other studies to utilize noise reduction for speech recognition. Lastly, for test device to be properly utilized by both patients and doctors, the tests must be in accordance to the usual tests done by ophthalmologists; the displayed results must be easy for the patient to understand; and the saved results must be what the ophthalmologists are expecting to see so that further analysis of the data is possible.

To address the points above, this study’s main objective is to create a system that tests three visual parameters: visual acuity, contrast sensitivity and color vision with the implementation of speech recognition for visual acuity and contrast sensitivity testing. The specific objectives are as follows: (1) To create a locally made system that can perform visual acuity test, contrast sensitivity test and color vision test; (2) to implement speech recognition for the of visual acuity and contrast sensitivity testing; and (3) to test the system and verify the results to actual results from a specialist.

With the high prevalence for visual impairment, and the short-handedness of ophthalmologists at certain areas, a device that can perform preliminary vision screening without the presence of the ophthalmologist but can provide results for the ophthalmologists can refer to in their diagnoses is of importance. More so since people who have visual impairments could encounter complications in their professional activities as well as risk their everyday lives. The significance of this study extends beyond the integration of multiple vision tests in one device. The tests that were chosen were for the visual parameters that could bring signs of eye diseases so that early detection and prevention is possible. The device can also enable doctors to test patients beyond their clinics, especially those in the rural areas, and easily interpret the results because the tests administered by the device conform with the usual tests done in the clinic. Cutting of equipment cost will allow different stakeholders to obtain and privately or publicly use the device.

The research only covers the detection of three visual impairments namely- blurry vision, low contrast sensitivity and colorblindness. There will only be three tests performed, visual acuity test, contrast sensitivity test and colorblindness test. Only the visual acuity and contrast sensitivity tests done using speech inputs, the colorblindness test. As such, the system may not be able to test patients with difficulty in speech for their visual acuity and contrast sensitivity. Patients who are incapable of following instructions like children below certain ages as well as the mentally challenged cannot be tested either. The researchers relied on the words of an ophthalmologist and related literature about the relation of visual acuity, contrast sensitivity and color vision to eye diseases and thus, will not actively search for samples with such eye diseases to become part of our sample. Finally, this study does not aim to replace ophthalmologist and is only for preliminary testing; should results show a cause for concern regarding possible eye diseases, it would be in the users’ discretion whether to contact eye doctors for a more detailed check-up.

**Chapter 2**

# **REVIEW OF RELATED LITERATURE**

## **The Different Visual Parameters**

### *Visual Acuity*

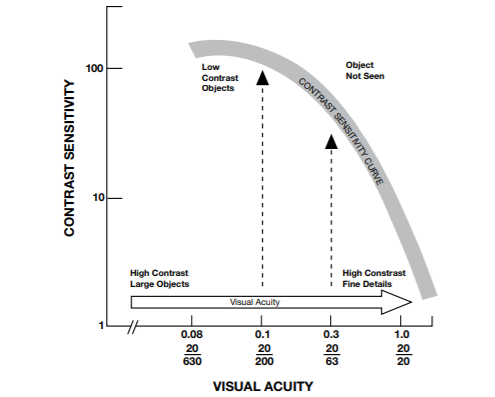
Based on its definition on Merriam-Webster, visual acuity is: “the relative ability of the visual organ to resolve detail that is usually expressed as the reciprocal of the minimum angular separation in minutes of two lines just resolvable as separate and that forms in the average human eye an angle of one minute” [14]. Refractive errors on eyes worsen one’s visual acuity. The WHO’s take on low vision defines it to be a visual acuity that is worse or less than 6/18 (20/60) but equal to or better than 3/60 (20/400) in the better eye when using the best correction possible [2]. With that Snellen’s equivalent, the person cannot properly view anything without his/her corrective lenses; everything is practically a blur. This poses occupational and safety risks at the times when the person does not use the prescribed corrective lenses. From the meta-analysis conducted by Huang et al. [15], the prevalence of refractive errors, more specifically myopia, is increasing not only in Asia, but also in the West. Another study by Pan et al. in 2012 concluded that as children, myopia may be more prevalent in Asia, more particularly with those possessing Chinese ethnicity; however, as adults, there are only slightly higher than that of the adults in the West [16]. It is starting to get common for people to develop refractive errors that when undergoing tests about a person’s visual function, visual acuity is one of the most commonly tested parameters in clinical practice [17].

### *Contrast Sensitivity*

The visual ability to distinguish objects from their backgrounds is known as contrast sensitivity. The objects in our field of vision vary in size, color, brightness; some objects are far, some are close; Some objects have myriad details while some are monotone. The images that enter the pupil in the form of light waves are sensed by the pupil, these images are then sensed by the retina and sent to that brain for interpretation. The difference in luminance of the object and its surroundings is knows as contrast. Increasing the contrast on your computer screen allows black to become darker and white to become whiter, mid-range colors start to disappear. The contrast sensitivity function “is a measure of contrast thresholds for a range of object sizes” [3].

A person’s contrast sensitivity can mean a presence of certain eye diseases. Galucoma, cataracts and diabetic eye diseases are known to reduce a person’s contrast sensitivity. With increasing cataract severity, the visual acuity of the patients decreased and there was a significant loss if contrast sensitivity. Contrast sensitivity was reduced significantly at high spatial frequency (18 cpd) in cortical cataracts with glare in day light and at low spatial frequency (3 cpd) in night light [18].

Contrast sensitivity curve or Visiogram can show the faintest contrast perceived by a person. Contrast sensitivity is plotted in the y-axis and visual acuity in the x-axis. The symbols decrease in size along the x-axis and becomes paler in the y-axis as shown in figure 2.1 below [19].



**Figure 2.1**. Visiogram

Usually contrast is expressed as percent, the ratio is then multiplied by 100. 100% is the maximum contrast. The symbols of the visual acuity charts are close to the maximum contrast. The sensitivity is determined by dividing 100 by the percentage of the lowest contrast perceived [19].

### *Color Vision*

Color vision is the ability of detect and distinguish colors. The human eyes possess special retinal cells known as cone cells that have different sensitivities to varying wavelengths of light; particularly for human vision, the 380mm - 750mm range is the one that is visible and distinguishable. The light's intensity is also a factor of how people perceive the colors surrounding them [20].

Color-vision deficiency (CDV), dyschromatopsia, or color blindness is a condition that limits a person by rendering him unable to differentiate some colors against the others to some degree [21]. A study by Lee and dos Santos listed the different classifications of colorblindness; it is said that having trichromatic vision means that the person can sense all three colors. However, there are anomalous trichromacy where the detection for one color is less sensitive than the detection of other colors; dichromatism which is the absence of a cone; and monochromatism is the total inability to sense color. Anomalous trichromacy can either be Protanomaly, Deuteranomaly or Tritanomaly while Dichromatism may be Protanopia, Deuteranopia or Tritanopia. The suffixes -anomaly and -anopia stand for “less sensitivity to” and “absence of” respectively. Much like the study by Lee and dos Santos, to easily and conveniently identify deficiency in color vision, the paper henceforth will identify the deficiencies by the root words *Protan, Deutan,* and *Tritan* to indicate red, green and blue deficiencies respectively [22].The prevalence of color blindness 6.5% of the population of China for male and 1.7% for females. A study on 2010 by Cruz, E.M. et al. deemed that the Philippines needed to monitor the people to determine if they are colorblind or not, the study focused on the prevalence of color-vision deficiency among Filipino male high-school students from Arellano High School. The study found that among the male students of the public high school, 5.17% (65/1258) of them are color blind with the majority being a deutan type deficiency [21].

## **Tests for the Different Visual Parameters**

### *Refractive Errors*

One component of having a good vision depends on visual acuity. Eye charts are used to determine the clarity of a person’s vision at a distance. One example of an eye chart is the 10 Early Treatment of Diabetic Retinopathy Study (ETDRS), which is said to be the most commonly used for visual acuity assessment within a research setting and is also the largest (10 item) well standardized set [23].

One of the scoring factors for ETDRS depends on how well a patient can distinguish a certain letter. Patients have the capability of being familiar to letters, however some patients are also unfamiliar with the Latin alphabet [24].The solution for this type of problem is by using an orientation judgement type of eye chart wherein patients are asked about the position rather than distinguishing the actual letter. Landolt C or also known as Landolt broken ring is an eye chart which features only one optotype, and the patient’s goal is to locate the position of the gap in a broken-ring stimulus [25]. Landolt C has been approved as the standard for visual acuity testing and has been used to evaluate contrast sensitivity [24].

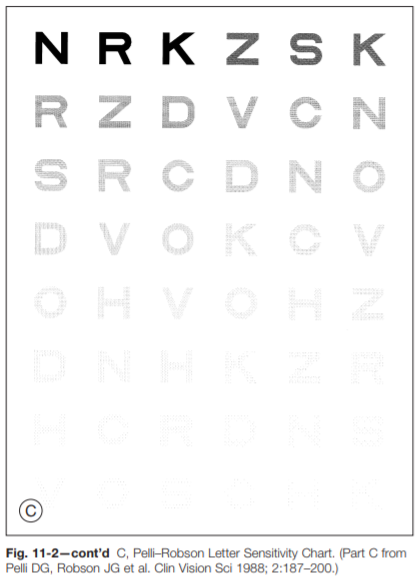
A study on 2011 by Becker, R., et al compared Landolt C and ETDRS charts for measuring visual acuity for both healthy patients and patients with various eye diseases. The result showed that there were only small and statistically not significant differences between Landolt C and ETDRS acuity [26].

T-test is a test that is utilized to compare the means of two values. This test is a type of parametric method which can be used when samples satisfy the conditions of normality, equal variance and independence.

A previous research by Lazaro et al. in 2018 has performed the T-test and was done with 10 samples. This research has a similar concept with the previous research which makes this test valid and reliable [13].

### *Contrast Sensitivity*

The testing is like measuring visual acuity at high contrast. We measure the smallest size of the optotype a person recognizes. Recognizing 3 out of 5 optotypes is the threshold. A widely used device for contrast sensitivity testing is the Pelli-Robson sensitivity chart as shown on Figure 2.2 below. This chart consists of horizontal lines of capital letters. The contrast in each successive line decreases, instead of the letters [27]. The most practical test used is the 2.5% test since the resulting threshold is far enough from the high contrast value so the declination of the slope in the curve can be defined. Figure 2.3 shows examples of 2.5% low contrast number and symbol charts. To record the results, the number of optotypes must be noted if one of the symbols on the 2.5% chart was incorrectly read on line 20/63 (6/18, 0.3) record the visual acuity value as 20/63 (-1) at 2.5% [19].



**Figure 2.2.** Pelli-Robson Sensitivity Chart. (G. Rubin, 2012 [27])



**Figure 2.3.** 2.5% Low Contrast Number and Symbol Charts

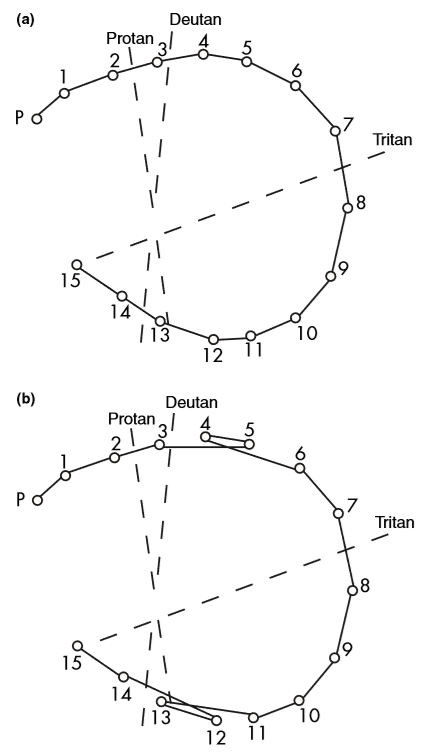
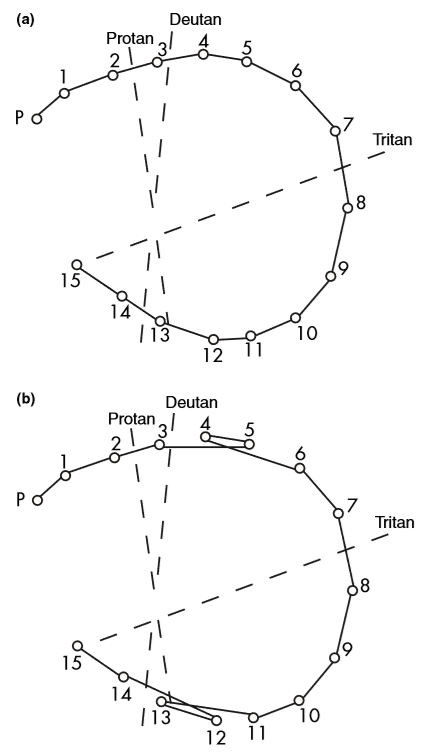
### *Color Blindness*

The current clinical vision tests available are generally divided into two categories: Pseudoisochromatic plate tests, and the Hue discrimination tests. Color vision tests such as Ishihara plates or the Farnsworth-Munsell 100-Hue test (FM-100) are usually performed by eye care practitioners during classical orthoptic assessment. They may sometimes be administered in order to determine a person’s ability to carry out a specific occupational task. These tests have been widely studied for a long time and their reliability; reproducibility and accuracy have been proven [28]. Computerized tests are often presented as a promising alternative to classical tests: They appear to be less expensive, allowing for automated treatment and storage and therefore less time consuming [4]. They can also be administered remotely, online where a variety of web sites propose digital versions of classical tests as Ishihara plates, or Farnsworth pawns. In a more rigorous approach, computerized tests have been validated by clinical trials and comparison with classical tests. For computerized testing, the Farnsworth-Munsell Panel D-15 (FMD-15) is often used due to its more rapid assessment; it is a shortened version of the FM-100 [28]. Dr. Glen V. Guevara, 2018, an ophthalmologist, suggested the use of FMD-15 over Ishihara plates since the test is capable in the examination of tritan defects.

Rather than the exact measurement of the severity in color blindness, the FMD-15 only classifies the type of color blindness that patients possess. The test specifically uses confusion axes to identify whether the problem is under protan, deutan or tritan [29]. FMD-15 consists of 16 colored caps: 15 being movable and the other is set as the reference cap. The caps, when a person with normal color vision arranges the D-15, are ordered side by side after the reference (P) based on the ascending numerical values of the hues shown on figure 2.4 [30]. The order of how people arrange such caps would determine the type of color blindness a person is suffering from or if the person does not have color blindness at all [29], [31]. A study in 2008 by Birch included the Figure 2.5a below that shows the correct order of how the FMD-15 colors should be arranged and the confusion axes of the three types of errors. The pass criteria include the correct order of hues or, at most, two single transformation of adjacent hue as shown in figure 2.5b. Classification of the isochromatic confusions would only be viable if there are at least two parallel crossings; furthermore, the point where strong deficiency is considered will be after ten crossings [29].



**Figure 2.4.** Farnsworth D-15 Hues in Their Correct Order. (L. Hyvärinen, n.d.)

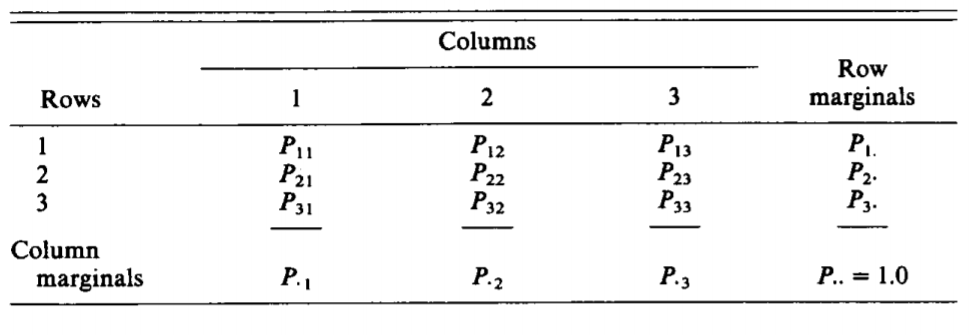
 

**Figure 2.5.** Test Result Diagrams: (a) Correct Order of Color Hues and (b) Two Single Transformation of Adjacent Hues. (J. Birch, 2008 [29])

The study by Birch in 2008 stated that, ideally, the sample for FMD-15 should contain equal number of protanopes, protanomalous trichromats and deuteranopes and ﬁve times this number of deuteranomalous trichromats. That is so that it would coincide with the general population’s prevalence. The study however did not include normal samples since the subjects were chosen from Ishihara test failures, as well as tritans because there were none who made major crossings in line with the tritan axis [29]. This lack of major crossing is due to the very low prevalence of tritan type defects which only has a prevalence of 1 in 13,000 worldwide [32]. From different researches about color vision tests, hundreds of samples were tested [29], [33], [34]. However, there are also, studies with lesser number of samples, particularly 18 and 25 samples [10], [35]. In choosing statistical tests for most medical research, an article by Nayak and Hazra in 2011 stated that Cohen’s Kappa statistic is best used when determining if there is statistical agreement between data sets if the data is categorical [36]. A study by Brennan and Prediger in 1981 about the Cohen’s Kappa showed the use of the kappa as a descriptive statistic that summarizes the agreement of two judges across several objects. Supposing two judges classify N samples into n categories, an n x n contingency table is formed as shown on table 2.1 [37]. The coefficient kappa by Cohen in 1960 is mathematically computed using equation 2.2 and is interpreted as the proportion of agreement after “chance” proportion of agreement is discounted: a kappa value of greater than 0.8 is deemed almost perfect in agreement [37], [38].

**Table 2.1**

3 x 3 Contingency Table (R.L. Brennan & D.J. Prediger, 1981 [37])



(2.2)

where:

= the observed proportion of agreement

= the “chance” proportion of agreement

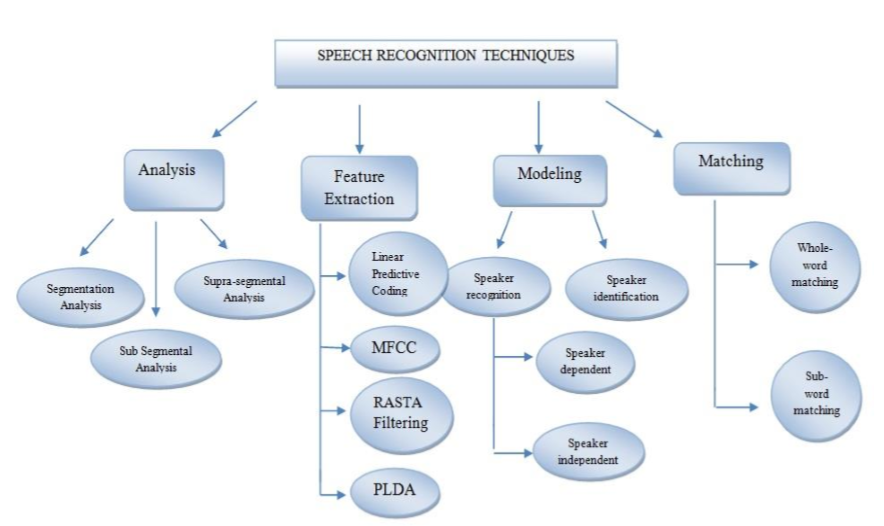
## **Computer-based Detection of Multiple Visual Impairments**

There have been multiple studies about the detection of visual impairment with the aid of computers, some of which are used to test single ocular condition while others are capable of different tests in a single program. A study done by Paglinawan et al. in 2017 proposed a raspberry pi-based detection of visual impairments so that people from the rural area who have little to no access to ophthalmologists can still get tested. Visual impairments including Strabismus via Hirschberg test; blind spots using visual field test; and blurry vision using visual acuity test are available in their system. The test results were stored in the local database of the Raspberry Pi so that they can retrieve the results in the future [11].

Another study about computer-based visual impairment detection was done by Huang, Ropelato and Menozzi back in 2015. The authors proposed a fully automatic and computerized self-vision-screening system which requires no examiner. The screening system included visual acuity, color vision via Ishihara plates, stereo vision and contrast vision in three shades. They have simplified the testing by manipulating the tests so that the only thing the patient will be asked to do throughout the four tests will be to identify which side of the four directions (i.e. up, down, left or right), do the Landolt-C ring open. The examination thus, became faster and more efficient. The system has been compared to corresponding apparatus that are often used for testing and have shown corroborative result [10].

## **Speech Recognition**

Speech recognition is a technology wherein the computer recognizes the input given via speech rather than using a keyboard [39]. It must have the ability to listen, analyze, extract, characterize, and then act on the spoken information. As mentioned by Narang (2015) et al, there are four stages in speech recognition which includes the Analysis, Feature Extraction, Modeling, and Matching which is shown in figure 2.6 [40]. Speech that are converted to texts automatically saves human efforts and most especially time in the transcription process. Health care and smart home technologies that are speech recognition based are of great help for the aged population and people with disabilities [41].

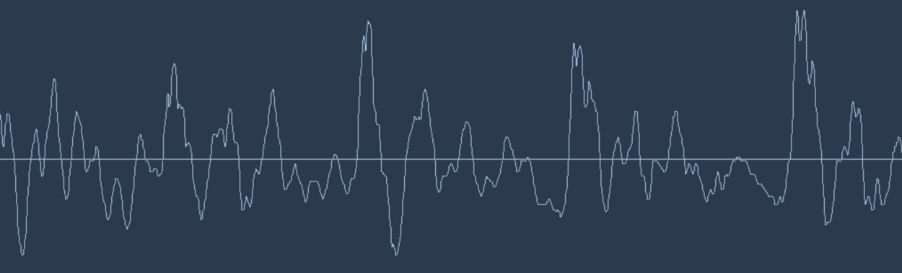


**Figure 2.6.** Speech Recognition Technique (S. Narang, 2015 [40])

All sounds created by human passes through the vocal tract which acts as a resonator and a filter. The shape of an individual’s vocal tract varies from one another hence the sound produced carries distinctive voice pattern [42]. The first stage is the speech analysis technique. This identifies the suitable frame size in segmenting the input speech signal and it creates raw features relating the envelope of the power spectrum of short speech intervals. Mainly, this stage is to analyze and characterize the speech. One analysis technique is the segmentation analysis in which it used in speaker recognition to extract the information due to vocal tract. The speech is evaluated using the frame size and shift in the range 10-30 ms [40], [43], [44].

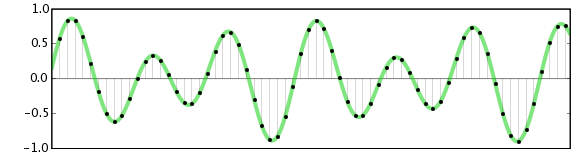
### *Input Speech*

The first step in speech recognition is to feed sound waves into computer. Sound is transmitted as waves which are one-dimensional (Fig.2.7) meaning at every moment of time, sound waves have a single value that is based on its amplitude [13].



**Figure 2.7.** One-Dimensional Sound Wave

To convert sound waves into numbers, which is something discrete, signal sampling is applied (Fig.2.8) at the Nyquist frequency. It permits a discrete sequence of samples to capture all the information of a speech signal. After sampling, the number that represents the amplitude of the wave at that time is recorded. The array of numbers gathered after sampling (speech data) is then pre-processed [13].



**Figure 2.8.** Sampled Sound Wave

The input speech signal includes noise. Noise are undesirable signals that interferes with the communication process hence it necessary to be removed. As seen on Fig 3.17 the input for the feature extraction is the input speech signal. Based on previous studies, feature extraction is the most critical stage since it extracts features that plays an important role for higher recognition accuracy, therefore, the noise from the input speech signal must be filtered [13].

### *Noise Removal*

The spectral subtraction method is a well-known single channel noise reduction technique. The technique applies subtraction of the noise spectrum estimate over the speech spectrum. It is also assumed that the noise component is relatively stationary [45]. The frequencies gathered in the speech signal produces a complex frequency since human speech is a complex sound. To make this data easier for process, a mathematical operation called Fast Fourier transform is performed. It is used to separate the complex and simple sound in order for the size of the data to be reduced. Performance of the process is faster with smaller training size. According to a study by Kumar in 2018, FFT's importance derives from the fact that in signal processing and image processing it has made working in frequency domain equally computationally feasible as working in temporal or spatial domain [46].

### *Feature Extraction*

Feature extraction is the most significant and critical stage in speech recognition. Extraction of features from the input speech signal plays an important role for higher recognition accuracy. There are various techniques that can be applied for feature extraction and each has its own pros and cons.

According to the study of Shanbhogue et.al in 2016, MFCC or Mel Frequency Cepstral Coefficients is the most dominant method that is used in speech recognition systems. MFCC can represent the low frequency region more accurately than the high frequency region and hence, it can capture formants which lie in the low frequency range and which characterize the vocal tract resonance [44]. The shape of the vocal tract displays in the envelope of the short time power spectrum, and in the method MFCC, the sound is represented as a short-term power spectrum on a non-linear Mel Scale of frequency [47]. Narang et al (2015) concluded in their study “Speech Feature Extraction Techniques: A Review”, that MFCC is the most used technique since it can imitate the human auditory system thus, gives a better performance rate [40]. 39 dimensional MFCC was used because a research comparing the 12, 26, and 39 dimensional MFCC as a feature extraction algorithm, it showed that with 39-dimensional MFCC, keyword spotting is more efficient and accurate [43].

As frequency gets higher, the main concern is how much energy occur at each spot. The Mel scale computes the distance needed on how to space filter banks and how wide to make them. The Mel scale relates perceived frequency of the speech signal to its actual measured frequency. Incorporating this scale makes the features match more closely to what humans hear. Equation 2.3 shows the formula for converting frequency to the Mel scale while equation 2.4 shows the formula for converting from the Mel scale back to frequency [48].

(2.3)

(2.4)

MFCC feature extraction starts by framing the signal into shorter frames.For each frame calculate the [periodogram estimate](http://en.wikipedia.org/wiki/Periodogram) of the power spectrum.Apply the mel filter bank to the power spectra, sum the energy in each filter. Summing up periodogram gives researcher an idea how much energy exist in various frequency regions. The Mel scale tells us exactly how to space our filter banks for efficient processing.Take the logarithm of all filter bank energies.Take the DCT of the log filter bank energies.Keep DCT 39 coefficients, discard the rest [49].

### *Feature Matching*

Dynamic Time Warping is an algorithm used for the comparison purpose for measuring similarity between two sequences which may vary in time or speed. Words are compared with reference word in DTW technique [40]. As a word is pronounced at different pitch, volume, speed etc. so every reference word has set of spectra, but no difference is observed between separate sounds in the word. Variable length of input signal and variance in speaking rate is handled with the help of DWT because it computes the similarity in two sequences that may vary in speed and time. Time warping is used to normalize the timing difference between sequences [43]. DTW is one of the algorithms used for measuring the similarity between two temporal sequences, which may vary in speeds. This method is used for Feature Matching.

### *Statistical Model*

Set of statistical models representing various sounds of the language to be recognized is the main core of all speech recognition. Hidden Markov model or HMM gives natural framework for constructing models that has a sequence of spectral vectors spanning the audio frequency range [50]. As mentioned by Rabiner et. Al, 95% word accuracy of speaker-independent task is achievable by using Hidden Markov model systems. It was also mentioned in his study that Hidden Markov model has a statistical methodology that is consistent and provides straightforward solutions to related problems and it provides inherent flexibility in various sophisticated tasks [51]. HMM is also used in lot of speech recognition system such as isolated word detection and continuous speech recognition because of its high reliability [51].

**Chapter 3**

# **METHODOLOGY**

## **Conceptual Framework**

**OUTPUT**

* Snellen Fraction for both visual acuity and contrast sensitivity.
* Saves circular test result to database, display number of major crossings and type of color blindness.

**PROCESS**

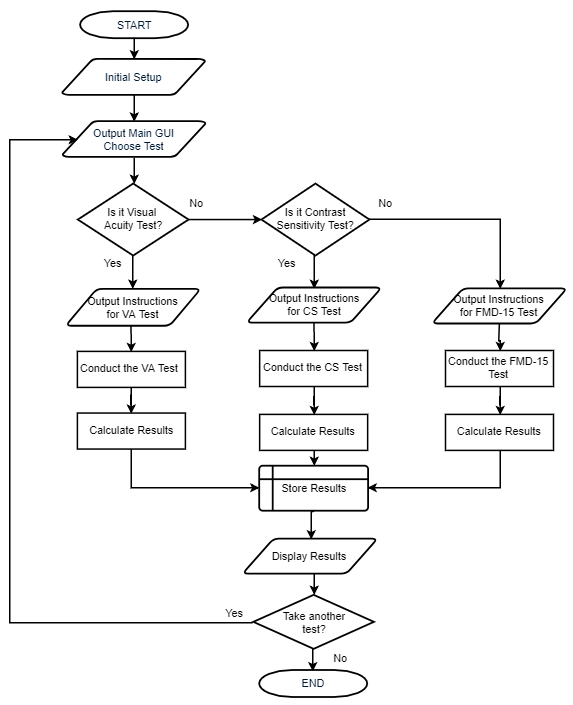
* Implementation of Speech Recognition.
* Plot a circular test result for Farnsworth D-15 test and count number of major crossings.

**INPUT**

* Patient Details
* Speech inputs for answers of patients
* Arranged Farnsworth D-15 hues for test using mouse.

**Figure 3.1.** Conceptual Framework

## **System Flowchart**



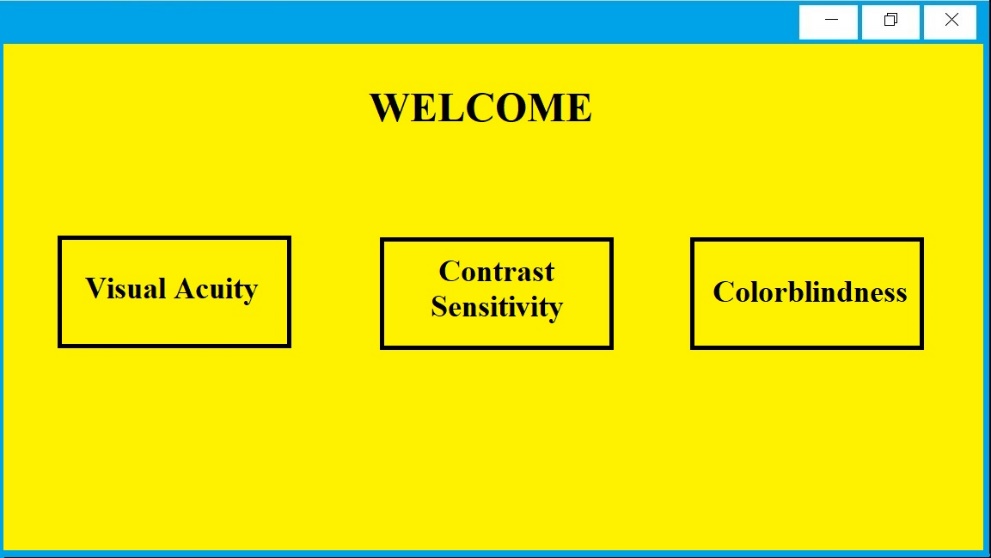
**Figure 3.2.** System Flowchart

The figure 3.2 shows the basic system flowchart of the device. After the initial setup i.e. booting up the Rpi and running the program, the main GUI will be displayed. There, the user may choose among the three different tests that are available which are the visual acuity test in the form of 4 direction Landolt-C optotypes; the 2.5% contrast sensitivity test also in the form of 4 direction Landolt-C optotypes; and the Farnsworth Munsell D-15 hue arrangement test for the assessment of color vision. Instructions are then provided by the device depending on the test chosen. After the instruction prompt, the test will be conducted in accordance to the methods which will be discussed in the following parts of the paper individually. The tests will be a combination of the outputs from the system and the inputs from the patient undergoing the said test including the personal details required for storage of data and the answers the patient will provide for the tests. The inputs from the patient vary depending on which test is chosen. For both visual acuity and contrast vision, a microphone is used as the input device for their speech answer regarding which orientation the Landolt-C optotype is facing while for the FMD-15, the mouse is used to arrange the colors by closest related hues.

After the chosen test is accomplished, the results are calculated by the device based on the chosen or constructed algorithms. These results will then be stored into the device’s local memory which is the micro SD card. After storing the result, it is then displayed in a chosen form for the patients to see and take note of. What to do with the knowledge of the results would then be in the patients’ discretion whether to seek for further in depth eye examination of an ophthalmologist or not, a copy of the results however, will be saved in case the owner of the device needs such data. With the results already out, one of two choices are made for whether another test will be done or not: If yes, a button is clicked and system returns to the main GUI; if no, the window is closed.

## **Objectives**

### *To create a locally made system that can perform visual acuity test, contrast sensitivity test and color vision test*



**Figure 3.3.** Proposed Main GUI

The proposed content of the Main GUI is shown in fig 3.3. Python will be used to code the graphical user interface. The library PyQt4 is installed in the Raspberry-Pi and will be utilized since we are able to design our own interface. This is a user-friendly interface is made to make this project easier to use. A data base is utilized to store the patients’ details such as name, birthday, age, location, medical history etc. SQLITE is utilized as a data base.

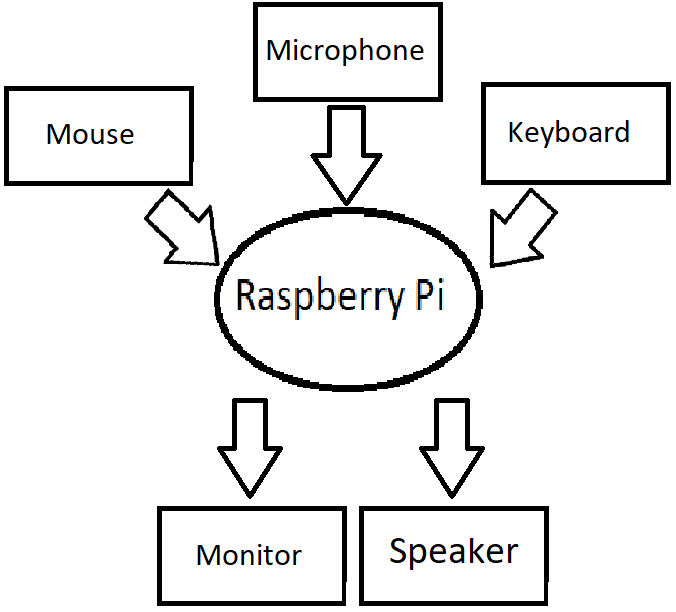
**SETUP**

The Raspberry Pi is programmed to perform visual acuity test, contrast sensitivity test ,and colorblindness test. Python will be used as the primary programming language.

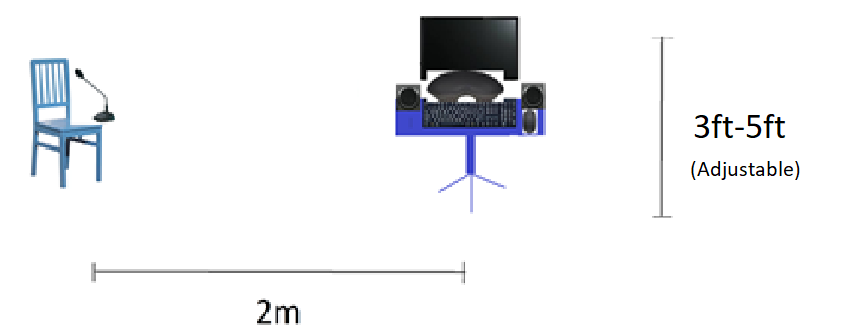
For the implementation of visual acuity test and contrast sensitivity test, the Yanmai SF-500B will be utilized as the microphone since it has a frequency range of 100Hz-16kHz which is perfect for voice recording and is the input device and for the implementation of the Farnsworth D-15 test, the mouse will be utilized as input devices.

The visual outputs will be dislayed on a 17 inch monitor, the acer V176L will be utilized since it has a maximum resolution of 1280 x 1024 which will be able to display optotypes clearly. The monitor is connected to a raspberry pi with the use of a HDMI to VGA adaptor. Speakers/heaphones may be utilized to output instructions. A block diagram of the input and output devices along with the microcontroller is shown in Figure 3.4

The setup will then be placed in an adjustable holder, in which the monitor can be adjusted at the eye level of the patient. Figure 3.5 shows the overview of the system’s setup.



**Figure 3.4.** Block Diagram of the Input and Output Devices Along with the Microcontroller



**Figure 3.5.** Overview of the proposed setup

**Visual Acuity Test**

Person

Speech

Speech Recognition

Comparison

Text

Score

Snellen Fraction

Vision

Acuity

Researchers

**Figure 3.6.** Flowchart on Visual Acuity Testing

**Visual Acuity Test**

The visual acuity of a patient can be described using three different scaling systems, this includes decimal visual acuity, Snellen fraction, and logMAR acuity. In this study, Snellen fraction will be used to describe visual acuity. Snellen fraction is expressed as the ratio of the test distance to the distance at which the critical detail of the smallest optotype resolved would subtend 1 minute of visual angle (or arc). This fraction is mathematically computed using equation 3.1.

(Equation 3.1)

Where:

= the visual acuity, measured as Snellen fraction;

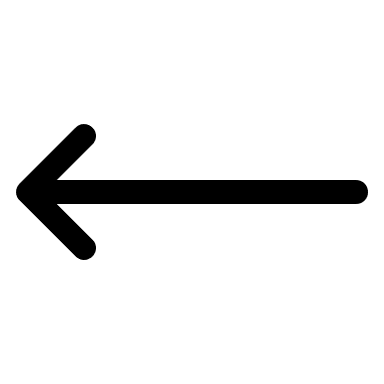
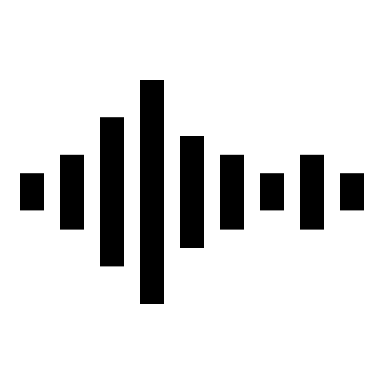
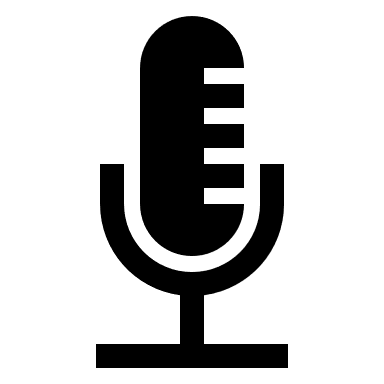
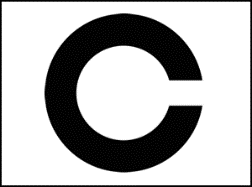
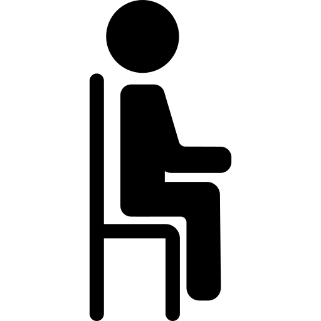
= the test distance, measured in m or ft;

= the normal distance, measured in m or ft.

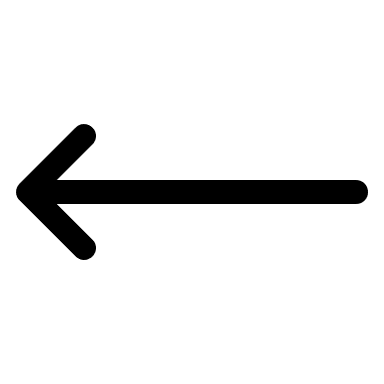
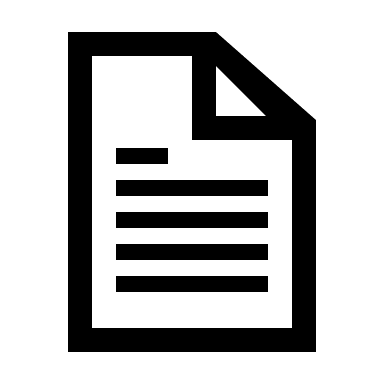
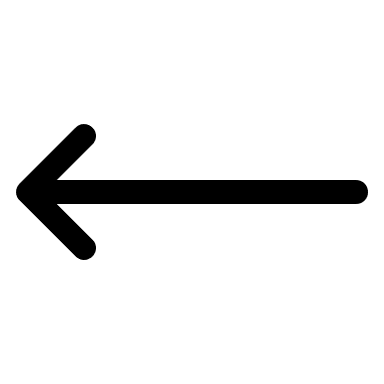
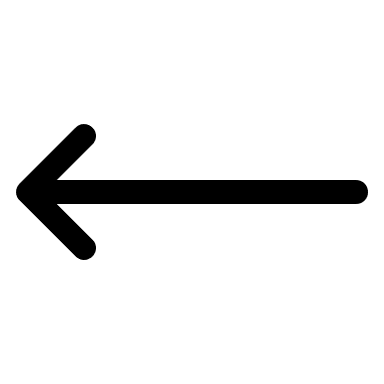
The landolt ring that will be used will have 4 different gap orientations which includes up, down, right, and left. In this study, 5 Landolt C that have the same sizes will be displayed in the monitor one at a time.

Figure 3.7 shows the schematic setup for the visual acuity testing. Each patient will be asked to provide trainer input commands to be saved in the system. Their voice will be recorded using a microphone uttering the words ‘up’, ‘down’, ‘right’, and ‘left’. The input speech responses are matched with the reference trained templates. Their speech records are added to the database or the reference trained input. After recording their speech, the patient is instructed to sit at 2 m away from the monitor and wear a test frame. The test frame is used to cover one eye since the vision is tested monocularly. When a Landolt-C appears on the screen, the subject identifies the orientation of the gap (Figure 3.8). On each trial, the Landolt-C is presented one at a time on the screen at one of four random orientation. The subject responds by speaking into a microphone, he/she is only allowed to speak four specific words:

* Up - if the gap is located at the top of the ring.
* Down - if the gap is located at the bottom of the ring.
* Right - if the gap is located at the right side of the ring.
* Left - if the gap is located at the left side of the ring.



Speech Recognition Process

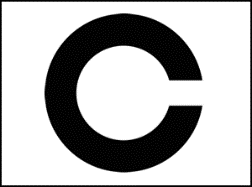
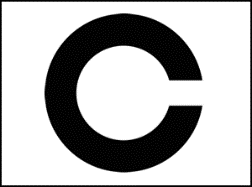
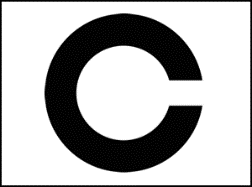
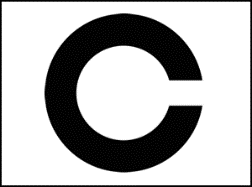


Up Down Right Left

Up Down Right Left

2 Meters

**Figure 3.7.** Schematic Setup for Visual Acuity Testing



RIGHT

LEFT

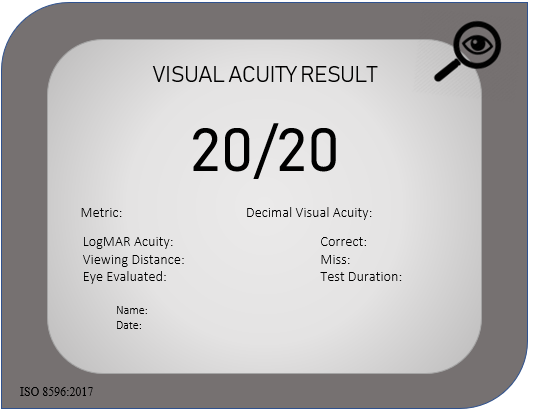
DOWN

UP

**Figure 3.8.** Landolt Rings at four orientations

Python is used to utilize the program for the speech recognition. After the input speech has been spotted in the reference trained templates, the system gathers the spotted keyword data for the analyzation of the eye test result.

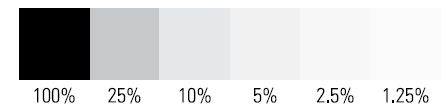
The sequencing of the presentation of the Landolt C is to present 5 targets at each size which will be displayed at the monitor one at a time. The response of the subject on each landolt ‘C’ will be scored dichotomously as correct (1) or incorrect (0). After each level, the response will be evaluated by checking if the total correct response is 60% of the total Landolt ‘C’. Since 5 Landolt ‘C’ are used, 3 correct responses are needed to have a “Passed” result. If the patient’s response has at least 3 correct responses, it will then proceed to the next smaller set of Landolt-C and so on. The test will be terminated at the first acuity grade for which the number of correct responses falls below 3. The visual acuity grade will be assigned as one grade lower than that at which the test was terminated. Figure 3.9 shows the test result for visual acuity (proposed GUI) which is shown after the test is completed.



**Figure 3.9.** Test Result for Visual Acuity (Proposed GUI)

**Contrast sensitivity**

Measurement of contrast sensitivity using low contrast visual charts will be utilized. The contrast sensitivity test is like the visual acuity test. We measure the smallest optotype the person can recognize. This test will be using the 2.5% test since this is the most practical test in clinical use. The resulting threshold point on the curve is far enough from the high contrast value so that the declination of the slope in the curve can be defined. Figure 3.10 shows 6 different contrast levels.



**Figure 3.10.** 6 Different Contrast Levels

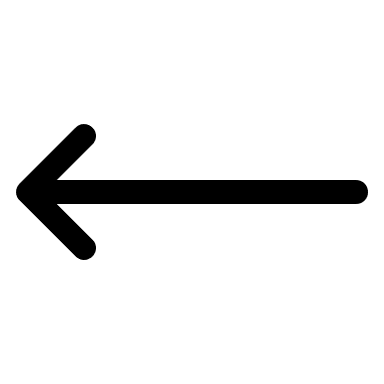
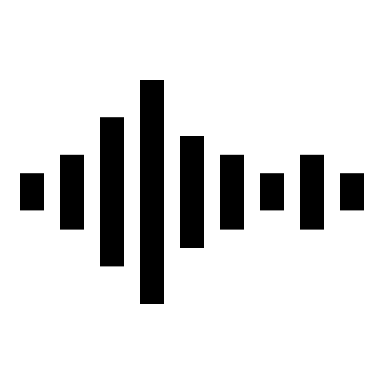
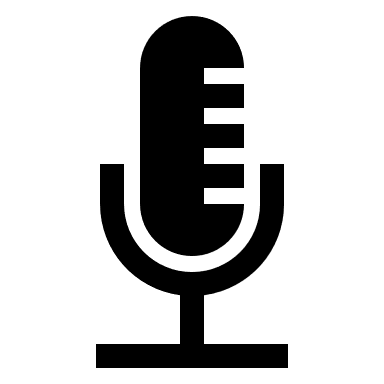
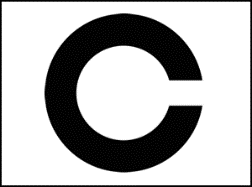
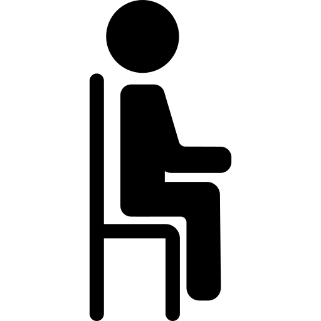
The optotype used will be Landolt C. The test will be conducted in the same manner as the visual acuity test, the only difference is that it will be done at 2.5% contrast. 5 Landolt C will appear on the screen one at a time per Snellen fraction. The inputs of the patients will then be compared to the actual answer and if the patient is able to get at least 60% of the characters correct since there will be 5 characters, the patient must at least get 3 out of 5 correctly for the next set of ‘C’ will appear. These ‘C’ will be smaller than the previous set of ‘C’. If the patient fails to get at least 60% of the characters correctly the test will terminate, and the result taken will be the previous Snellen fraction.

The patient will be asked to sit 2 meters away from the monitor. The patient will then follow a set of audio instructions in order to perform the test. The patient is to respond “up”, ‘down’, “left” and “right” to indicate the direction of the broken ring. This will be done on both eyes, one at a time. Figure 3.11 shows the optotypes and the correct responses. Figure 3.12 shows the setup of the test.

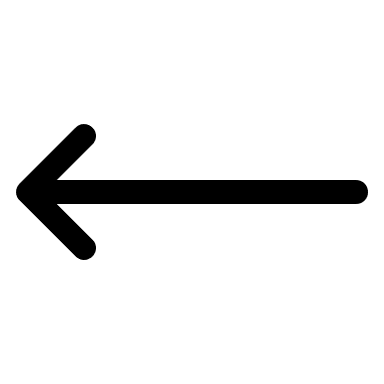
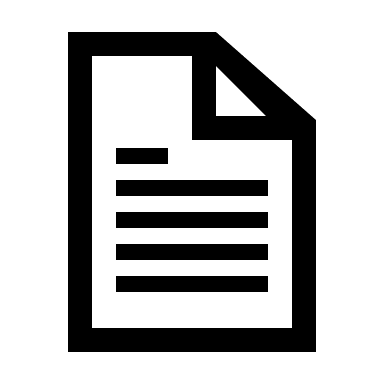
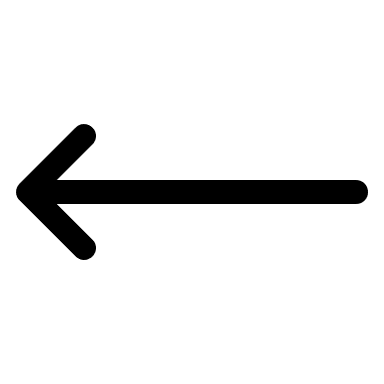
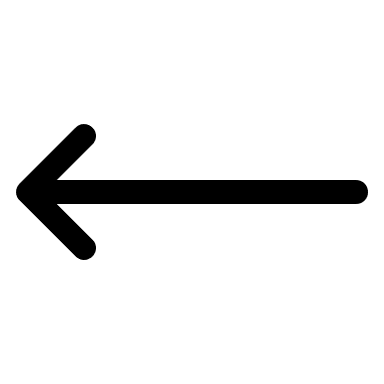


Right Down Left Up

**Figure 3.11.** Possible Optotypes and Respective Correct Answers



Speech Recognition Process



Up Down Right Left

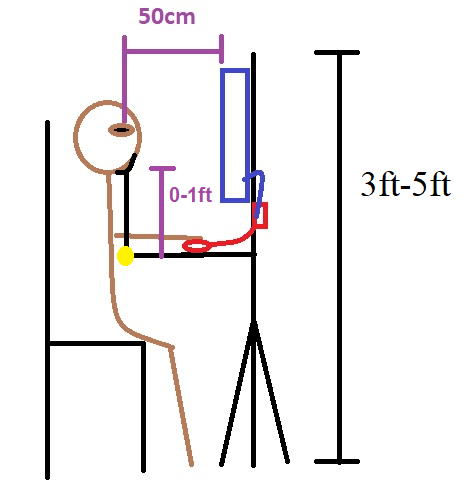
Up Down Right Left

2 Meters

**Figure 3.12.** Schematic Setup for Contrast Sensitivity Testing

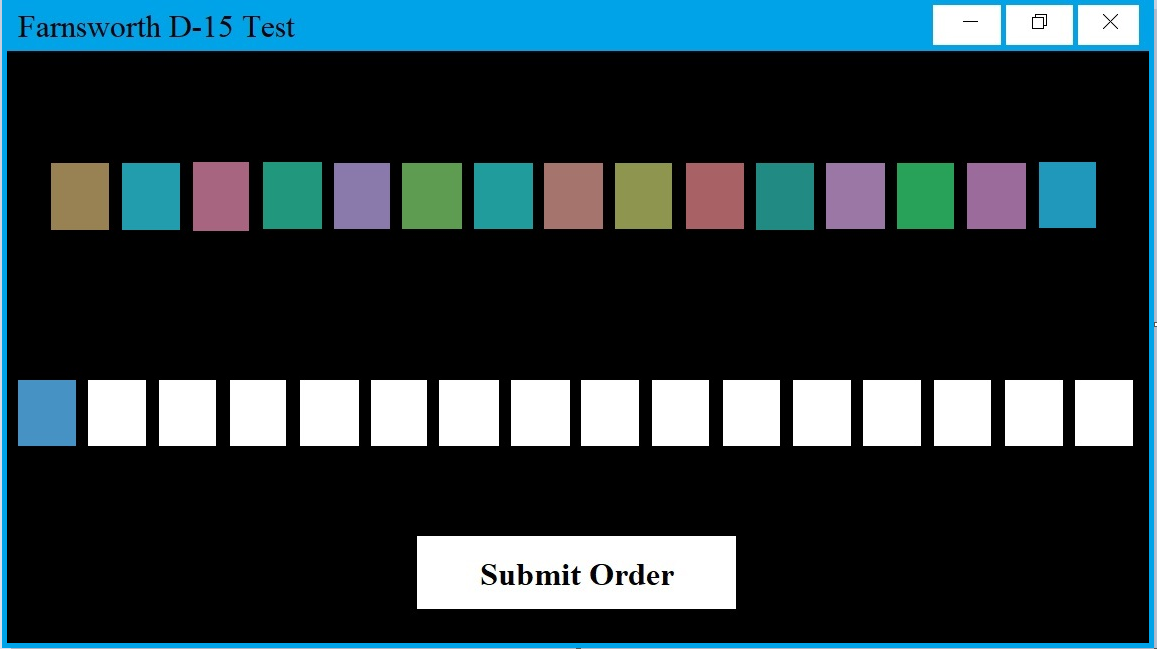
**Fransworth D-15 test**

For the color blindness test, the FMD-15 is used in order to classify the type of color blindness the patient has. This is because FMD-15 can classify the 3 types of deficiencies and it boasts fast testing time. Using both eyes, the patient is positioned at 50cm from the monitor and the monitor is positioned at eye level as shown on figure 3.13. A chinrest is made adjustable by 0 to 1 foot in height so that the patient will be able to take the test at a fixed position from the monitor 50cm away.

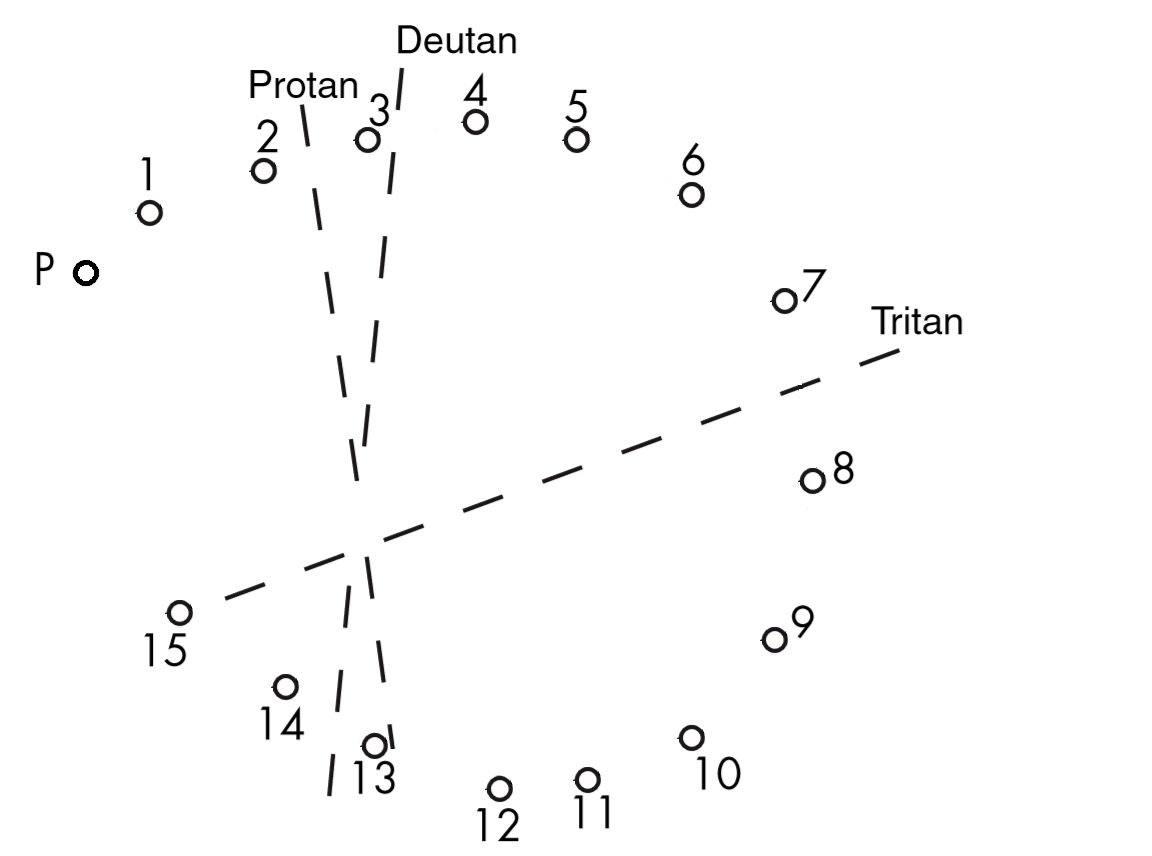


**Figure 3.13.** FMD-15 Setup

The hues are arranged randomly in a line as shown in the test’s user interface in figure 3.14. Based on the figure, the test is administered with black background for maximum contrast in order to avoid wrong perception of color due to surroundings. The test is done by rearranging the colors by hues beside the reference hue. The user must order the hues according to how close the hues are with each other by placing them in the blank, white spaces. Hues may be reordered prior to submission of order. After the submission of the test, the order that the hues are arranged are recorded and compared with the normal order and the confusion axes. Figure 3.15 shows the test result template as well as the confusion axes where the circles will be connected based on the patient’s ordered hues.

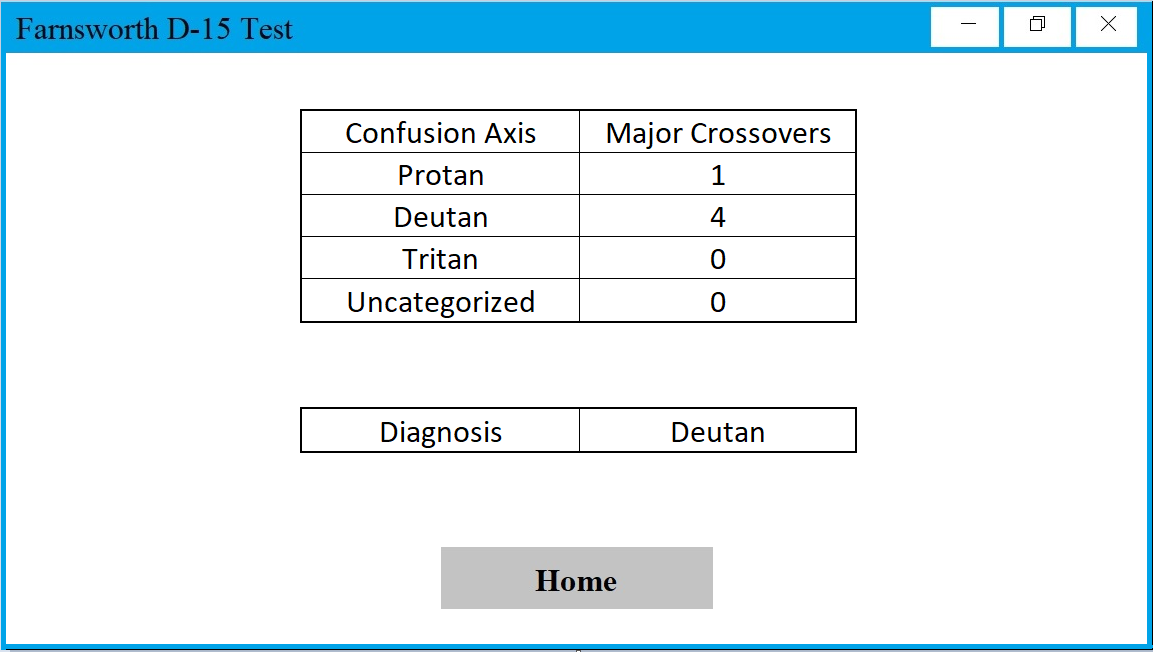


**Figure 3.14.** FMD-15 test user interface

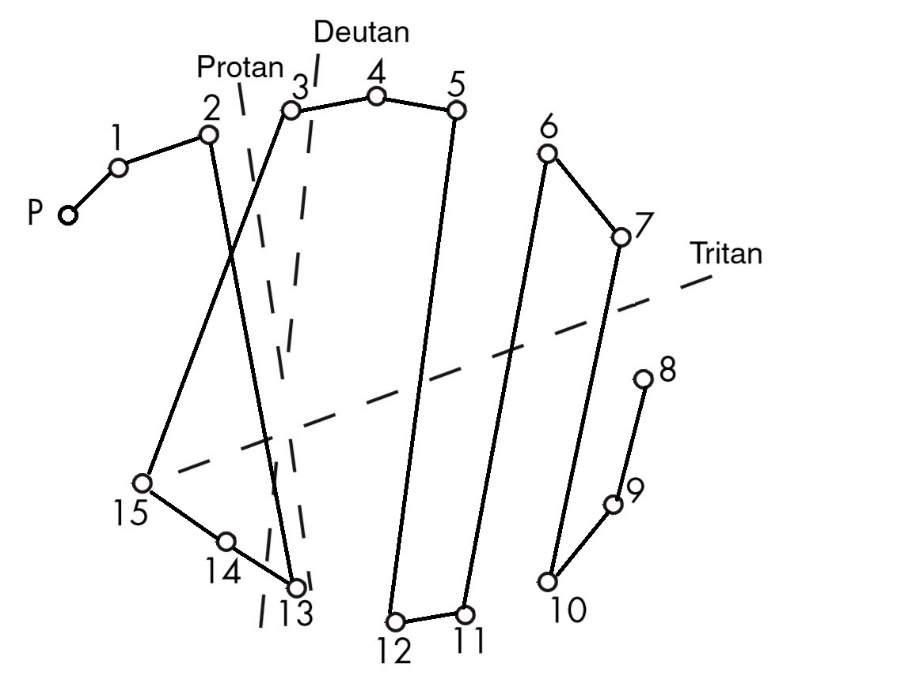


**Figure 3.15.** FMD-15 test result template with confusion axes

Since FMD-15 can be defeated by multiple trials when patients receive results via the numbered caps, only results that count the number of major crossings as well as the diagnosis will be shown after the test as shown on figure 3.16. However, the test results that are stored, shown on figure 3.17, will display which major crossings are made using the scoring template of the FDM-15 so that further analysis of results can be made by doctors.



**Figure 3.16.** Color Blindness Test Result (to be Displayed Only)



**Figure 3.17.** Color Blindness Test Result (to be Stored in the Database Only)

**Calibration and Controlled Tests**

For visual acuity and contrast sensitivity, we will be displaying the optotypes in a 17-inch monitor, the chart needs to be properly calibrated for it not to affect the results. We first should calibrate the size of the optotype with respect to the distance of the patient from the screen using the equation 3.2. An arcminute is a unit of angular measurement that is equivalent to 1/60 of a degree.

H=D\*tan() (3.2)

Where: H = Height of the letter in millimeter

D=Distance from the screen

After adjusting the first line the researches will proceed to adjusting the next lines.

For controlled testing of the system for both visual acuity and contrast sensitivity, forced errors to monitor false results were introduced. The researchers intentionally made the responses incorrect to check if the system is working as it is intended to. We will also have samples of known visual acuity and contrast sensitivity and test them to check if our system is calibrated.

For calibration purposes, the researchers introduced forced response testing of the FMD-15. Table 3.1 shows the different forced response made to the system in the form of the displayed (e.g. Figure 3.16) and stored (e.g. Figure 3.17) test results, as well as the interpretation of the system and the expected interpretation for such a test. A total of five forced responses are made to ensure the different diagnosis of the results are viable.

**Table 3.1.**

Results for Forced Response Testing for FMD-15

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Controlled Test # | Displayed Result | Stored Result | System’s Interpretation | Expected Interpretation |
| 1 | Figure 3.5 | Figure 3.6 | Deutan | Deutan |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| 5 |  |  |  |  |

### *To implement speech recognition for the of visual acuity and contrast sensitivity testing.*

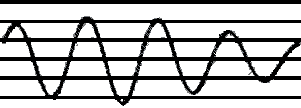
Input Acoustic Frame Sequences

Speech



Dictionary of Phonemes

Sound Wave Pattern



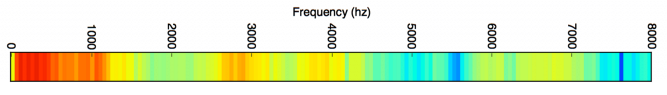
Hidden Markov Mode

Recognized Speech

Acoustic Frame Sequences

Frequency Spectogram

Recognized Speech



Acoustic Frame Sequences

**Figure 3.18.** Recognition Process

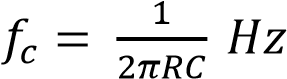
The process for recognizing the uttered speech as an input for the system is shown in figure 3.18. Speech recognition process starts by capturing the sound uttered in a microphone. Sound is transmitted as waves which are one-dimensional, meaning at every moment of time, it has a single value based on amplitude. The sound wave is converted into something discrete and signal sampling is applied. Fourier transform, which is a mathematical operation, is applied to the extracted digital signal. It breaks down sound waves into complex and simple sound. These energies contained in each individual sound waves are added to get an idea of how much energy exists in those various frequency regions. After that, it is then converted to its spectrogram data representation.

The next step is the fragmentation of digital data into sequence of corresponding acoustic frames. The data is then evaluated to identify the components of speech they compromise. Features are then extracted and is associated to the predefined dictionary of phonemes. The phonetic dictionary predefined by the researchers are composed of the words’ fragmented sounds, list of scripts, numbers, commands.

1. **Band Pass Filter design and Noise Removal**

Frequency range of 300 Hz - 3000 Hz is where all speech activities lies. Therefore, the system only need the frequencies in that definite range. This range of frequency is the acoustic frames or energy being referred to. For the system to reject signal component outside that range, band pass filter is implemented.

Equation 3.3 is calculation for the cutoff frequency which can be applied for both low-pass filter and high-pass filter needed for the implementation of the band pass filter. Resistance as well as the capacitance value can also be determined.

 (3.3)

**Noise Removal**

After the band pass filter is implemented, noise removal system takes place. For the noise removal block, the system removes the noise from the input speech signal using Spectral Subtraction.

**Steps:**

1. Speech Input Partitioned into Frames / Buffers with a predefined amount of overlap.

2. Initial Noisy Frames are Windowed, Transformed to Frequency Domain (FFT) and Averaged

3. Each Frame is Windowed, Transformed to Frequency Domain (FFT) and Subtracted by Average Noise.

4. Half Wave Rectification is performed

5. Reconstruction through IFFT and Overlap-Add Method

1. **Energy Sum and Threshold**

For the system to identify if a word has been uttered by a subject, energy threshold is specified. Energy accumulator acts as a low-pass filter and is gathered at a slow rate so that high frequency background sound that will occur at a short period of time while speech recording can be filtered This allows the microcomputer to recognize the utterance of the subject without any interruption. The microcomputer begins storing samples for further processing if the energy reaches a specified threshold. Energy threshold is calculated using the equation 3.4 below.

(𝑛) = (𝑛) + (1 − 𝛼)(𝑛 − 1) = 𝑦(𝑛 − 1) + 𝛼(𝑥(𝑛) − 𝑦(𝑛 − 1)) (3.4)

Equation for Energy Threshold

Where: y(n) = new value for energy

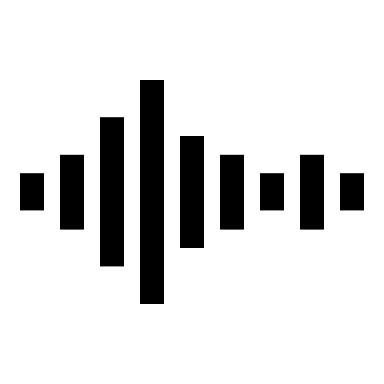
x(n) = absolute value of the difference between the input signal from the microphone and the DC offset provided by the microphone

𝛼 = time constant of the energy accumulator

After the subject has spoken and the energy threshold is reached, the microcomputer will start collecting samples. Collecting samples only takes place after the energy threshold is crossed and before energy threshold starts to decay.

1. **Logical Structure**



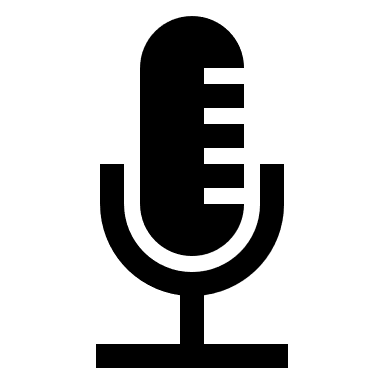


FFT on Speech Sample

User’s Speech Sample

Microphone

System



Keyword Spotting

Print Result or Continue Testing

Comparison of Indices to Dictionary

Hidden Markov Model

**Figure 3.19.** Logical Structure of Speech Recognition

Figure 19 above shows logical structure on how speech recognition is used in performing the visual acuity and the contrast sensitivity test. The test starts with the monitor displaying the largest Landolt-C optotype down to the smallest one where the test has been terminated due to the wrong responses of the user. The subject speaks into the microphone then it is inputted to the raspberry pi microcomputer obtaining the sample rate of the speech sample. The raspberry pi microcomputer further processed the sampled rate by storing the first and last samples into two separate arrays. Fast Fourier Transform will then be applied on the samples for both separate arrays to transform it to frequency domain. With the use of HMM or Hidden Markov Model, the dictionary created by the programmers and the sorted bins with their corresponding indices are then compared to identify which keyword in the dictionary most closely matched the speech input. This will be the basis if the result of the eye test will be printed or if it needs to display another Landolt-C optotype to continue and repeat the process to gather more information from the subject.

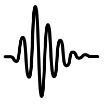
1. **Hidden Markov Model (HMM)**

As mentioned by Mohan and Babu (2014), Hidden Markov Models (HMM) has been extensively used in a lot of speech recognition applications such as isolated word detection and continuous speech recognition due to its high reliability [52]. Hidden markov models are usually used in speech tagging, is it to determine hidden parameters from an observable data Speech recognition system has compromises of a set of statistical models demonstrating several sounds to be recognized. HMM creates stochastic models from known utterances and compares the probability that the unknown utterance was generated by each model. The natural framework for the set of statistical models is provided by the HMM.

The steps for the application of HMM:

Given the observation sequence **O** and a model **λ,**

1. Make Sound arrangement represented by variable S for demonstrating, for example, phonemes or words, then call the sound classes 𝑉 = {𝑣1, 𝑣2, … , 𝑣𝑠).
2. In individual class, gather a sizable set, the preparation set, of considered noises that are identified to be in the class.
3. Depending on every preparation set, compute the estimation problem to acquire the best model **λ** for each class 𝑉𝑖 (𝑖 = 1, 2, … , 𝑆).
4. While doing the assessment, evaluate Pr(**O**|**λi)** (𝑖 = 1, 2, … , 𝑆) the unfamiliar utterance **O** and classify every speech which formed O as class 𝑉𝑖 .
5. **HMM Based Recognizer**



Speech Waveform

Feature Extraction

Acoustic Models

**Y**

Feature Vectors

Pronunciation Dictionary

Decoder

Language Model

Words

“Start”

**W**

**Figure 3.20** Architecture of an HMM Based Recognizer

Figure 3.20 shows the principal components of a speech recognizer. The task for a recognizer in speech recognition is given an unknown signal, the uttered word must be determined. An audio waveform from the microphone is converted to a fixed size of acoustic vector via feature extraction method. The acoustic vectors extracted is represented by 𝑌1:𝑇 = 𝑦1, … . , 𝑦𝑇. The function of the decoder is to look for the indices or sequence of words characterized by 𝑤1:𝐿 = 𝑤, … . , 𝑤𝐿 which is most likely to have generated **Y.** The decoder block attempts to find w using equation 3.5.

𝑤 = 𝑎𝑟𝑔𝑚𝑎𝑥 {𝑃 (𝑤|𝒀)} (3.5)

Difficulty is encountered in modeling {𝑃 (𝑤|𝒀)} directly. For this problem to be resolved, Baye’s Rule is implemented. Baye’s Rule transforms eq. 3.5 into the equivalent problem of identifying w in equation 3.6.

𝑤 = 𝑎𝑟𝑔𝑚𝑎𝑥 {𝑃 (𝒀|𝑤)(𝑤)} (3.6)

The likelihood p(Y |w) is determined by an acoustic model and the P(w) is identified by the language model.

The acoustic model is represented by phones since it is a basic unit of sound. The acoustic model of any given word is created by concatenating the phones which it is made up of. These are concatenated to make words as defined in the pronunciation dictionary. N-gram model is used as the language model. The probability of each word is conditioned only on its N-1 predecessors and its parameter are estimated in a text corpus by counting N-tuples. The decoder then operates by searching through all possible word sequences. To eliminate improbable searches and to decrease the time for decoding, pruning is used.

1. **Feature Extraction**

This stage extract features on the speech waveform that can give a good match made by the acoustic models. Mel Frequency Cepstral Coefficients or MFCC was used for feature extraction. MFCC can represent the low frequency region more accurately than the high frequency region and hence, it can capture formants which lie in the low frequency range which characterize the vocal tract resonance. Sounds generated by human speech are filtered by the shape of the vocal tract. And that shape manifests itself in the envelope of the short time power spectrum, and the purpose of MFCC is to accurately represent that envelope.

**Steps:**

1. Frame the signal into shorter frames. The signal is framed into 20-40 ms in this process.
2. For each frame calculate the [periodogram estimate](http://en.wikipedia.org/wiki/Periodogram) of the power spectrum.
3. Apply the mel filterbank to the power spectra, sum the energy in each filter.
4. Take the logarithm of all filterbank energies.
5. Take the DCT of the log filterbank energies.
6. Keep DCT coefficients, discard the rest.
7. **HMM Acoustic models**

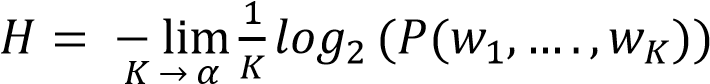
HMM is exhibited to three characteristics model. These three characteristics are the observation distribution, model topology, and decoding hierarchy. The spoken word *w* is fragmented into a sequence (𝐾𝑤) of base phones. The sequence represents its pronunciation 𝑞1:𝐾𝑤 = 𝑞1, … . . , 𝑞1:𝐾𝑤. Each word can be pronounced differently, hence the possibility of having multiple pronunciations must be considered. To allow multiple pronunciations, p (𝒀|𝑤) can be calculated using equation 3.7. Only valid pronunciation of sequences for 𝑤 is included in the summation.

p (𝒀|𝑤) = ∑𝑄 (𝒀|𝑸)(𝑸|𝒘), (3.7)

The conditioning of word history is necessary for large vocabulary recognition. It it is truncated to N − 1 words to form an N-gram language model where N is in the range 2-4.

 (3.8)

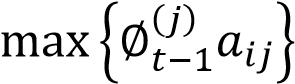
Language models are assessed in terms of their perplexity H. This approximation is intended for finite length word sequence which are used for modelling a N-gram. Perplexity (H) is defined in equation 3.9.

(3.9)

1. **Decoding**

This process is where sequence of hidden states is discovered given the sequence of observations. Given as input an HMM λ = (A,B) and a sequence of observations O= o1, … . , oT, the most probable sequence of states Q = q1q2q3 ...qT is to be determined . The Viterbi algorithm is commonly used for decoding and is now standard for application of dynamic programming. Viterbi algorithm computes for the maximum probability to look for the most probable sequence of hidden states that will become an outcome in sequenced of observed events or to simply put, it computes for the most probable sequence of states.

(3.10)

In using equation 3.10 above, the first step is to initialize , 1 is used for the initial, non-emitting, entry states, and 0 for all other states.  is the probability of the most probable word sequence. Recording the gathered data in every maximization decision yields a traceback of the required best matching state/word.

**Hardware Implementation**

* Raspberry Pi
* 4 Amp Power Adapter
* 64 GB Micro Sd
* Microphone - Installation of the PyAudio package is needed for the accessability of the microphone

### *To test the system and verify the results to actual results from a specialist.*

**How the Device Works:**

The patient is to be seated 2m from the setup for the visual acuity test and the patient is to follow voice instructions during the procedure.

**The instruction and commands from the speakers are:**

Cover your right eye.

* This command instructs the patient to cover his/her right eye for visual acuity and contrast sensitivity testing of the left eye

Cover your left eye.

* This command instructs the patient to cover his/her left eye for visual acuity and contrast sensitivity testing of the right eye

Begin arranging the colors.

* This command means the patient may start arranging the colors.

A tone is also played every time an input for the test is requested in visual acuity and contrast sensitivity tests. After the test is complete, the results will be shown.

**Visual Acuity Test Using Landolt-C Optotypes**

10 participants participated in the visual acuity test without correctors and were chosen using general advertisements or recruitment. Additional criteria include being able to perform the test without correctors and that the patients will have to take an acuity test from an optometrist for actual results if they have not reported a result from a test within the last 12 months. Table 3.2 shows the summary of results for the visual acuity tests per patient and it is filled as follows:

**Table 3.2.**

Results for Visual Acuity Test

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Patient’s Name: | | | | | | | Date: | | |
| Snellen Fraction | 1st C | 2ND C | 3RD C | 4TH C | 5TH C | Correct | Miss | Result |
|  |  |  |  |  |  |  |  |  |
| 0.08 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| 1.1 |  |  |  |  |  |  |  |  |

For the second column (1st C) up to sixth column (5th C):

* “ ” (blank)– if the response of the subject is correct
* “0” – if the response of the subject is incorrect

For the seventh column (Correct):

* Total number of “ ”(blank)– shows the total correct responses of the subject

For the eight column (Miss):

* Total number of “0” – shows the total incorrect responses of the subject

For the ninth column (Result):

* “Passed” – if the total number of correct responses at each size level is either 5, 4 or 3.
* “Failed” – if the total number of correct responses at each size level is either 2, 1 or 0.

The null hypothesis for the t-test is that the mean Snellen fractions obtained from the system and specialist is statistically be the same. The t-test will be utilized to compare the results. We would reject the null hypothesis of the test will have a p < 0.05 since this would mean the two means are not statistically the same. Table 3.3 shows the results of the system and the results given by the ophthalmologists.

H0 = (Null hypothesis): If p > 0.05, then it is statistically same.

H1 = (Alternative hypothesis): If p < 0.05, reject Null hypothesis.

**Table 3.3.**

Results for Visual Acuity Test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Patient** | **Left eye** | | **Right eye** | |
|  | System result | Specialist Result | System result | Specialist Result |
| A |  |  |  |  |
| B |  |  |  |  |
| C |  |  |  |  |
| D |  |  |  |  |
| E |  |  |  |  |
| F |  |  |  |  |
| G |  |  |  |  |
| H |  |  |  |  |
| I |  |  |  |  |
| J |  |  |  |  |

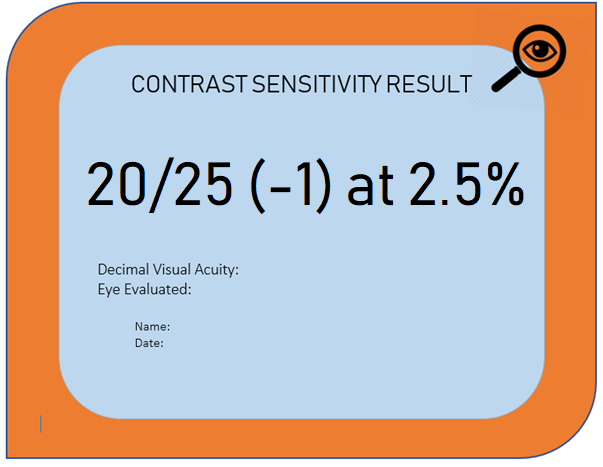
**Contrast Sensitivity Test**

10 participants participated in this test. These participants are the same as the one from the visual acuity test. Table 3.4 is filled up by noting the number of correct and incorrect responses from the patient. An ”x” will be placed at the nth c that the patient has an incorrect response. The number of correct answers will appear on the “correct” column and the number of missed symbols will be placed on the “missed column”. In the results column “passed” or “failed” will be the placed on the column. “Passed” would indicate that the patient got at least 3 out 5 of the symbols correctly and “failed” would indicate otherwise. If the number of missed symbols at the Snellen fraction exceeds two, the previous Snellen fraction will be shown as the output else the result will be shown as “Snellen Fraction(-M)” such that M is the number of misses. Figure 3.21 shows a sample GUI of the output.

**Table 3.4.**

Results of Contrast Sensitivity Test at 2.5%

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Patient’s Name: | | | | | | | Date: | | |
| Snellen Fraction | 1st C | 2ND C | 3RD C | 4TH C | 5TH C | Correct | Miss | Result |
|  |  |  |  |  |  |  |  |  |
| 0.12 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| 0 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| 0 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| 1.60 |  |  |  |  |  |  |  |  |
| 2.00 |  |  |  |  |  |  |  |  |



**Figure 3.21.** Test Result for Contrast Sensitivity (Sample GUI)

For statistical verification of the contrast sensitivity test, much like visual acuity test, t-test will be used. Table 3.5 shows the results of the system and the results given by the ophthalmologists.

**Table 3.5.**

Results for Visual Acuity Test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Patient** | **Left eye** | | **Right eye** | |
|  | System result | Specialist Result | System result | Specialist Result |
| A |  |  |  |  |
| B |  |  |  |  |
| C |  |  |  |  |
| D |  |  |  |  |
| E |  |  |  |  |
| F |  |  |  |  |

**Color Blindness Test using Farnsworth Munsell D-15**

The system’s tests for the FMD-15 have been done in accordance with the settings and procedures above. To test the device with actual patients, the researchers took 20 samples for testing through general advertisements and personal recruitment. The sample patients consisted to 2 with normal color vision, 4 protans, 12 deutans and 2 tritans. Number of samples for protans and deutans are at a ratio of 1:3 because of combining Birch’s anomalous trichromacy and -anopes (i.e. 1+1 for protans and 5+1 for deutans), however, this study will include both samples with normal color vision as well as tritans in order to check the capabilities of the system in correctly identifying normal color vision and tritan defects. The number of tritans, however, is less than the others due to their rarity. Two trials for each sample are done as to avoid the samples in becoming overly familiar with the test. Since there are two trials made in succession, the results of the first trial (i.e. in the form of figure 3.16) that come after the test are hidden from the sample but will be captured by the researchers and will only be shown together with the results of the second trial (i.e. also in the form of figure 3.16) after both trials have been completed while the results (i.e. also in the form of figure 3.17) are stored and recorded for verification. These trials will be treated as separate samples and thus, the test will have a total of 40 samples. Table 3.6 below shows the summarized form of the actual testing. The summary consists of seven columns that include data such as Patient number; results for both trial 1 and 2 in the form of stored results; the diagnosis of the device; as well as the diagnosis of an ophthalmologist based on the results.

Cohen’s Kappa test will be used to statistically determine the agreement between the system’s result and the actual results. Four Kappas will be determined: one for each type of diagnosis each using all the 40 samples that are taken from a population of 20. Thus, for the actual results, there will be 4 with normal color vision, 8 protans, 24 deutans and 4 tritans. Table 3.7 records the contingency matrix for normal, protan, deutan and tritan. Equation 2.2 is used to calculate Cohen’s Kappa when there are two tests and a kappa greater than 0.6 means the tests are substantially in agreement and a kappa beyond 0.8 is deemed to be near perfect agreement. This study hypothesizes that the system is in near perfect agreement with the doctor’s diagnosis for normal color vision, protan, deutan and tritan. The hypotheses are as follows:

H0 = (Null hypothesis): the system results agree with ophthalmologists.

H1 = (Alternative hypothesis): the system results do not agree with the ophthalmologists.

**Table 3.6.**

Summary of Results for Color Vision Test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Patient number | Trial 1 | Trial 2 | Classification  (Device) | Classification (Ophthalmologist) |
| xxx | Fig CVxxx.1 | Fig CVxxx.2 | 1) Normal  2) Normal | Near normal |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| 5 |  |  |  |  |
| 6 |  |  |  |  |
| 7 |  |  |  |  |
| 8 |  |  |  |  |
| 9 |  |  |  |  |
| 10 |  |  |  |  |
| 11 |  |  |  |  |
| 12 |  |  |  |  |
| 13 |  |  |  |  |
| 14 |  |  |  |  |
| 15 |  |  |  |  |
| 16 |  |  |  |  |
| 17 |  |  |  |  |
| 18 |  |  |  |  |
| 19 |  |  |  |  |
| 20 |  |  |  |  |

**Table 3.7.**

4 x 4 Contingency Matrix

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| System Results | Actual Results | | | | |
| Normal | Protan | Deutan | Tritan | R. Marinals |
| Normal |  |  |  |  |  |
| Protan |  |  |  |  |  |
| Deutan |  |  |  |  |  |
| Tritan |  |  |  |  |  |
| C. Marginals |  |  |  |  | 1.0 |

# 

# **References**

[1] D. C. Van Essen, “In: Visual Neurosciences (L. Chalupa and J. Werner, eds.),” *Cortex*, no. Panel D, pp. 1–14, 2002.

[2] World Health Organization, “Priority eye diseases,” *Priority eye diseases*, 2018. .

[3] P. Bex, “Contrast Sensitivity,” *Ophthalmol. Clin.*, pp. 383–388, 2010.

[4] I. Benkhaled, I. Marc, and D. Lafon, “COLOUR CONTRAST DETECTION IN CHROMATICITY DIAGRAM : A NEW COMPUTERIZED COLOUR VISION TEST LGI2P IMT Mines Ales C2MA IMT Mines Ales,” pp. 1–4, 2017.

[5] National Eye Institute, “Facts About Diabetic Eye Disease,” 2015. .

[6] J. Van Eenwyk, A. Agah, and G. W. Cibis, “Automated human vision assessment using computer vision and artificial intelligence,” *2008 IEEE Int. Conf. Syst. Syst. Eng. SoSE 2008*, 2008.

[7] D. de Fez, M. J. Luque, L. Matea, D. P. Piñero, and V. J. Camps, “New iPAD-based test for the detection of color vision deficiencies,” *Graefe’s Arch. Clin. Exp. Ophthalmol.*, vol. 256, no. 12, pp. 2349–2360, 2018.

[8] B. C. Navitsky and S. Editor, “The Portable Eye,” no. December, pp. 15–17, 2015.

[9] J. B. Ibarra *et al.*, “Measurement of overall decentration, angle deviation, and prism diopters in categorized strabismus cases using mathematical morphology algorithm,” *HNICEM 2017 - 9th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag.*, vol. 2018–Janua, pp. 1–4, 2018.

[10] Y. Y. Huang, S. Ropelato, and M. Menozzi, “Fully Automatic and Computerized Self-Vision- Screening System,” pp. 262–266, 2015.

[11] A. C. Paglinawan *et al.*, “Detection of three visual impairments: Strabismus, blind spots, and blurry vision in rural areas using Raspberry PI by implementing hirschberg, visual field, and visual acuity tests,” *HNICEM 2017 - 9th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag.*, vol. 2018–Janua, pp. 1–9, 2018.

[12] N. G. Chua, “Vision 2020: Philippine Update 2005.”

[13] J. B. Lazaro, R. G. Garcia, A. L. Generalo, M. A. M. Halili, and M. S. Montebon, “Speech recognition for control of optotype characters of the Snellen chart using LogMAR transformation,” in *AIP Conference Proceedings*, 2018, vol. 2045, no. December, pp. 020055-1-020055-5.

[14] Merriam Webster Online, “Visual Acuity [Def. 1],” *Merriam Webster*. .

[15] J. Huang *et al.*, “Efficacy comparison of 16 interventions for myopia control in children: A network meta-analysis,” *Ophthalmology*, vol. 123, no. 4, pp. 697–708, 2016.

[16] C. W. Pan, D. Ramamurthy, and S. M. Saw, “Worldwide prevalence and risk factors for myopia,” *Ophthalmic Physiol. Opt.*, vol. 32, no. 1, pp. 3–16, 2012.

[17] A. Bastawrous *et al.*, “Development and validation of a smartphone-based visual acuity test (peek acuity) for clinical practice and Community-Based Fieldwork,” *JAMA Ophthalmol.*, vol. 133, no. 8, pp. 930–937, 2015.

[18] J. H. Shandiz *et al.*, “Effect of cataract type and severity on visual acuity and contrast sensitivity,” *J. Ophthalmic Vis. Res.*, vol. 6, no. 1, pp. 26–31, 2011.

[19] M. E. Hartnett, M. Trese, A. J. Capone, B. J. B. Keats, and S. M. Steidl, *Pediatric Retina*, First Edit. Philadelphia, USA: LIPPINCOTT WILLIAMS & WILKINS, 2004.

[20] D. Q. Utama, T. L. R. Mengko, R. Mengko, and M. N. Aulia, “Color blind test quantification using RGB primary color cluster,” *2016 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2016 - Proc.*, pp. 1–4, 2017.

[21] M. Emerson M. Cruz, M. Herma Grace S. Cerdana, M. Ann Margaret B. Cabrera, M. Chanda B. Garcia, M. Evelyn T. Santos-Morabe, and M. Ma. Lourdes R. Nañagas, “Prevalence of color vision deficiency among male high school students,” *Philipp J Ophthalmol*, vol. 36, no. 1, pp. 20–24, 2011.

[22] J. Lee and W. P. Dos Santos, “Fuzzy-based simulation of real color blindness,” *2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC’10*, pp. 6607–6610, 2010.

[23] L. M. Hamm, J. P. Yeoman, N. Anstice, and S. C. Dakin, “The Auckland Optotypes: An open-access pictogram set for measuring recognition acuity,” *J. Vis.*, vol. 18, no. 3, p. 13, 2018.

[24] K. R. Alexander and J. J. McAnany, “Determinants of contrast sensitivity for the tumbling e and landolt C,” *Optom. Vis. Sci.*, vol. 87, no. 1, pp. 28–36, 2010.

[25] M. V. Danilova and V. M. Bondarko, “Foveal contour interactions and crowding effects at the resolution limit of the visual system,” *J. Vis.*, vol. 7, no. 2, p. 25, 2007.

[26] J. S. Pointer, “Recognition versus resolution: A comparison of visual acuity results using two alternative test chart optotype,” *J. Optom.*, vol. 1, no. 2, pp. 65–70, 2008.

[27] G. S. Rubin, *Visual Acuity and Contrast Sensitivity*, Fourth Edi., vol. 1. Elsevier Inc., 2012.

[28] A. N. French, K. Rose, and E. Cornell, “The Evolution of Colour Vision Testing,” no. January 2008, 2008.

[29] J. Birch, “Pass rates for the Farnsworth D15 colour vision test,” *Ophthalmic Physiol. Opt.*, vol. 28, no. 3, pp. 259–264, 2008.

[30] L. Hyvärinen, “Colour Vision.” [Online]. Available: http://www.lea-test.fi/en/vistests/occupati/cv/occ\_cv.html. [Accessed: 22-Dec-2018].

[31] H. B. McBain *et al.*, “The impact of strabismus on quality of life in adults with and without diplopia: A systematic review,” *Surv. Ophthalmol.*, vol. 59, no. 2, pp. 185–191, 2014.

[32] B. L. Cole, “Assessment of inherited colour vision defects in clinical practice,” *Clin. Exp. Optom.*, vol. 90, no. 3, pp. 157–175, 2007.

[33] R. K. Shrestha and G. S. Shrestha, “Assessment of Color Vision Among School Children: A Comparative Study Between The Ishihara Test and The Farnsworth D-15 Test.,” *JNMA. J. Nepal Med. Assoc.*, vol. 53, no. 200, pp. 266–269, 2016.

[34] D. Oliphant and J. K. Hovis, “Comparison of the D-15 and City University (second) color vision tests,” *Vision Res.*, vol. 38, no. 21, pp. 3461–3465, 1998.

[35] J. S. Ng and S. C. Liem, “Can the Farnsworth D15 Color Vision Test Be Defeated through Practice?,” *Optom. Vis. Sci.*, vol. 95, no. 5, pp. 452–456, 2018.

[36] B. Nayak and A. Hazra, “How to choose the right statistical test?,” *Indian J. Ophthalmol.*, vol. 59, no. 2, p. 85, 2011.

[37] R. L. Brennan and D. J. Prediger, “Coefficient kappa: Some uses, misuses, and alternatives,” *Educ. Psychol. Meas.*, vol. 41, no. 3, pp. 687–699, 1981.

[38] J. Cohen, “A Coefficient of Agreement for Nominal Scales,” *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.

[39] D. Pandey, “Voice Recognition using VHDL,” *Int. Conf. Inven. Syst. Control*, pp. 1–4, 2017.

[40] S. Narang and M. D. Gupta, “Speech Feature Extraction Techniques : A Review,” *Int. J. Comput. Sci. Mob. Comput.*, vol. 4, no. 3, pp. 107–114, 2015.

[41] R. Sultana and R. Palit, “A Survey on Bengali Speech – to – Text Recognition Techniques,” *2014 9th Int. Forum Strateg. Technol.*, pp. 26–29, 2014.

[42] B. J. Baars and N. M. Gage, *Fundamentals of cognitive neuroscience: A beginners guide*. Amsterdam: Academic Press, 2013.

[43] K. A. Senthildevi and E. Chandra, “Keyword spotting system for Tamil isolated words using Multidimensional MFCC and DTW algorithm,” *2015 Int. Conf. Commun. Signal Process. ICCSP 2015*, pp. 550–554, 2015.

[44] A. VH and R. Marimuthu, “A Study on Speech Recognition Technology,” *J. Comput. Technol.*, vol. 3, no. 7, pp. 2278 – 3814, 2014.

[45] R. M. Udrea, S. Ciochina, and D. N. Vizireanu, “Reduction of Background Noise from affected Speech using a Spectral Subtraction Algorithm based on Masking Properties of the Human Ear,” *TELSIKS 2005 - 2005 uth Int. Conf. Telecommun. Mod. Cable Broadcast. Serv.*, vol. 1, pp. 135–138, 2005.

[46] T. S. Kumar, “FFT BASED CLOSED LOOP LINEAR SYSTEM RESPONSE WITH & WITHOUT DELAY ELEMENTS,” *Int. J. Curr. Eng. Sci. Res.*, vol. 5, no. 1, pp. 10–14, 2018.

[47] V. Singh, V. K. Jain, and N. Tripathi, “A Comparative Study on Feature Extraction Techniques for Language Identification,” *Int. J. Eng. Res. Gen. Sci.*, vol. 2, no. 3, pp. 286–291, 2014.

[48] R. I. A. Casil, E. D. Dimaunahan, M. E. C. Manamparan, B. G. Nia, and A. A. J. Beltran, “A DSP Based Vector Quantized Mel Frequency Cepstrum Coefficients for Speech Recognition,” *Inst. Electron. Eng. Philipp. J.*, vol. 3, no. 1, pp. 1–5, 2014.

[49] M. M. Oo, “Comparative Study of MFCC Feature with Different Machine Learning Techniques in Acoustic Scene Classification,” *Int. J. Res. Eng.*, vol. 5, no. 7, pp. 439–444, 2018.

[50] M. Gales and S. Young, “The Application of Hidden Markov Models in Speech Recognition,” *Found. Trends® Signal Process.*, vol. 1, no. 3, pp. 195–304, 2007.

[51] B. H. Juang and L. R. Rabiner, “Hidden Markov Models for Speech Recognition,” *Technometrics*, vol. 33, no. 3, pp. 251–272, 1991.

[52] B. J. Mohan and N. Ramesh Babu, “Speech recognition using MFCC and DTW,” in *2014 International Conference on Advances in Electrical Engineering, ICAEE 2014*, 2014, pp. 1–4.