Term Paper Report

on

# HUMAN BRAIN INTERACTION AND MOTION PERCEPTION USING BLUE EYES

Submitted to

Amity University
Uttar Pradesh

In partial fulfilment of the requirements for the award of the degree of

Bachelor of Technology in

Computer Science and Engineering by

**ASHISH KALRA**

**A2305219112**

Under the guidance of

**DR GARIMA AGGARWAL**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AMITY SCHOOL OF ENGINEERING AND TECHNOLOGY

AMITY UNIVERSITY, UTTAR PRADESH

MAY 2020

# DECLARATION

I Ashish Kalra student of B.Tech (2-CSE-2Y) hereby declare that the project title "**HUMAN BRAIN INTERACTION AND MOTION PERCEPTION USING BLUE EYES TECHNOLOGY**" which is submitted by me to Department of Computer science, **Amity School of Engineering Technology**, Amity University Uttar Pradesh, Noida, in partial fulfillment of requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering, has not been previously formed the basis for the award of any degree, diploma or other similar title or recognition.

The Author attests that permission has been obtained for the use of any copy righted material appearing in the Dissertation / Project report other than brief excerpts requiring only proper acknowledgement in scholarly writing and all such use is acknowledged.

Date: ___________

Ashish Kalra
A2305219112
2CSE-2Y (2019-23)

# CERTIFICATE

This is to certify that Mr Ashish Kalra, student of B.Tech in Computer Science and Engineering has carried out work presented in the project of the Term paper entitle "**HUMAN BRAIN INTERACTION AND MOTION PERCEPTION USING BLUE EYES TECHNOLOGY**" as a part of First year program of Bachelor of Technology in Computer Science and Engineering from Amity University, Uttar Pradesh, Noida under my supervision.

Dr Garima Aggarwal

Department of Computer Science and Technology

ASET, Noida

# ACKNOWLEDGEMENT

The satisfaction that accompanies that the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all efforts with success. I would like to thank Prof (Dr) Abhay Bansal, Head of Department-CSE, and Amity University for giving me the opportunity to undertake this project. I would like to thank my faculty guide Dr Garima Aggarwal who is the biggest driving force behind my successful completion of the project. She has been always there to solve any query of mine and also guide me in the right direction regarding the project. Without her help and inspiration, I would not have been able to complete the project. Also, I would like to thank my batchmates who guided me, helped me and gave ideas and motivation at each step.

Ashish Kalra

# ABSTRACT

The world has a herculean nature and now due to the increasing stress and workload in daily life, it has turned crueler than ever. In between this, the I.T giant IBM came up with a brilliant idea which focuses to make the computer wise enough to acknowledge human feelings and act accordingly as the emotional support and a worthy friend for their owner. Just imagine you turn on your computer and its greats with you a very warm welcome note and then just by a touch of you it understands that you want to read your e-mails or want to listen to that new music your favourite singer just dropped out.

The humans have perceptual abilities and that's why they can understand, express and share their feelings with others. This technology deals with providing the personal computer these perceptual abilities which include facial expressions, sensory touch and speech recognition to act as the intimate partner. This blue eye technology uses various elements of Artificial intelligence, Machine Learning and Data Analysis along with image processing abilities to curb the gap between physical and electronic world.

This jaw dropping technology has various advantages as it has reduced the overall fatigue of the user and has made various task seem to be easier than ever. The technology has step in foots into various modern devices including smartphones and smart television. The need for the sense enabled devices is increasing day by the day and so is the demand for the same. The internet connectivity and AI has made use of this technology spread to wide number of users across the world and provides each one of them with an intimate partner which understands them and act accordingly.

As most of the technologies its aim is to make the task easier but it is equally capable to protect and reduce the risk of various accidents and prominent damages and thus its implementation is far beyond the idea it was developed for and so it has become even more prominent in the industry.

# Contents

DECLARATION .... i
CERTIFICATE .... ii
ACKNOWLEDGEMENT .... iii
ABSTRACT .... iv
1. INTRODUCTION .... 1
  1.1. COINAGE OF THE TERM BLUE EYES .... 1
2. LITERATURE SURVEY .... 2
3. TECHNICAL OVERVIEW .... 3
  3.1. HARDWARE .... 3
  3.2. SOFTWARE .... 6
4. EMOTION COMPUTING .... 7
  4.1. THEORY .... 7
  4.2. RESULT OF THEORY .... 8
  4.3. TYPES OF EMOTION SENSORS .... 8
5. FOUNDATION OF GAZE TRACKING .... 10
  5.1. CLASSIFICATION OF EYE GAZE .... 10
  5.2. PRIMARY MODE OF GAZE ESTIMATION .... 10
  5.3. ESTIMATION OF EYE GAZE TRACKING .... 12
6. METHODOLY USED IN GAZE ESTIMATION .... 13
  6.1. 2D REGRESSION BASED MODEL .... 13
  6.2. 3D REGRESSION MODEL .... 14
  6.3. CROSS RATIO BASED METHODS .... 14
  6.4. APPEARANCE BASED METHOD .... 14
  6.5. SHAPE BASED METHOD .... 15
7. MAGIC POINTING .... 17
8. SUITOR .... 20
9. SPEECH RECOGINITION USING ARTFICIAL INTELLIGENCE .... 21
  9.1. HIDDEN MARKOV MODEL .... 21
  9.2. DYNAMIC PROGRAMMING .... 22
10. APPLICATIONS .... 24
11. CONCLUSION .... 25
12. REFERENCES .... 26

# 1. INTRODUCTION

The Blue Eyes technology is aimed to completely change the way humans interact with their personal computers. The idea behind this magnificent technology is to make computers use perceptual abilities integrated in them to interact with its user. The perceptual abilities will be used to identify user's identity, feel its presence and interpret his/her thoughts. The computers so made are equipped with most modern cameras, microphones and imparted sensual abilities to integrate and interpret audio- visuals and sensory abilities. Researchers have been able to impart more abilities like recognizing human presence, talks and feelings. Though their still lies problem regarding calibration of these abilities in order to suit the best to the environmental conditions.

The system can recognize using facial recognition, artificial intelligence and speech recognition to identify a user as it has the ability to gather information. The emotion mouse can feel your presence and understand all your emotions with just a touch.

It would be good to see a personal computer being as the intimate partner for a human and working as an emotional support in addition to just being a regular computer running day to day work-based errands.

Humans despite being quite skilled, educated and full of knowledge can make errors but the machines are often found to be perfect than their human counterparts. Even if a small percentage of human cognizing power is imparted to the personal computers, the computation would be much faster and accurate in terms of results. [1]

## 1.1. COINAGE OF THE TERM BLUE EYES

As name suggests Blue as in Blue Eyes stand for Bluetooth Technology which was developed by Dr Jaap Haartson while working at Ericsson in 1990's and the Eyes refer to the eye movements which are detected by the integrated cameras and provide lot of interesting intriguing information.

The Bluetooth technology allows with the easy dependable cordless communication and through private area for the networking and uncomplicated linking of the operators to the most vital central system.

The key features for this technology include:

1. Monitoring of visual and gaze attention
2. Monitoring of physiological condition
3. Detection of operator's position
4. Acquisition of wireless data over Bluetooth
5. Trigger real time user defined alarms
6. Recording of various parameters such as physiological condition, speech and gaze. [2]

# 2. LITERATURE SURVEY

The idea for the Blue Eyes Technology given by the IBM which is amongst the huge name in the IT world during the 1990's. The idea spread quickly as the need for computers that could interact with the humans seemed to be quite pleasing and came with known benefits to the mankind. Many developments were seen in the late 90's and 2000 which were planning to change the computing world. The idea of gaze estimation and MAGIC pointing came into play where the user could control his PC just by using his eyes later development for emotion sensing came into the lime light with the help of physiological detectors. [1] [3]

The idea started from the IBM's Almaden Research Institute team working in California during 1997 and later attracted many IT companies to look for the potential in the technology that could totally change the human brain interaction with the computers. Later Ericsson, the sole creator of the Bluetooth and Microsoft jumped in along with Poznan University of Technology.

The idea was to recognize human emotions and to use them for the benefits of the user by reducing workload, stress levels, providing an intimate partner and to understand user psychological condition. The paper by Benjamin M. [3] throws light on the Paul Ekman works on the various human emotions and performed physiological test on various participants and then combined the data of human emotions to it. He came up with six basic human emotions that humans are known to experience [1] . With the data compiled and calibrated the emotion sensing came into the light. More work was conducted in this area and now computers and other electronic gadgets can recognize human emotions. The paper by Benjamin M. [3] also talks about MAGIC pointing and its implementation and how it is suitably the best may instead of traditional manual inputs and the eye gaze. He explains the work done by Shumin Zhai in the 1997 and explains its practical implementation on a 25-inch screen.

Gaze estimation and interaction is most viable part for the Blue Eye technology and its algorithms have different approaches and methods to be performed to get the desired results. These algorithms and approaches along with their methodology are well explained in paper by Anuradha Kar [4] where she talks about various approaches and methodologies and explains how eye gaze estimation is different for different devices and has evolved over time. She has provided with various algorithms and gaze estimation methods along with emotion sensing from them.

V.V. Nanavare [5] talks about how emotion detection through speech and artificial intelligence have become quite prominent and helps to create a bond between the system and the user. He compares various algorithms for speech detection and explains how feature extraction works.

# 3. TECHNICAL OVERVIEW

Blue eyes system continuously monitors the operators physical and mental state i.e. saccadic activity along with operators with attention and head acceleration to keep a check on his cardiovascular and respiratory network. The system collects lethysmographic signals from forehead along with body temperature and evaluate the circulation rate and operator's oxygen levels.

It examines all the important levels and emotion data for any abnormality in users' body and warn the user by triggering pre-defined alarms as set by the user. The system checks values of blood oxygen levels and pulse rates. A user can interact with the system in case of a panic and the system will work on the operator's speech and act accordingly.

A mobile device with integrated Bluetooth is used which is connected to data sensors put upon by the user of the system and the most vital central unit wherein all data is analyzed. RFID enabled id cards are provided to various operators all linked to unique profiles to organize data from various operators who act as data source for long time analysis.

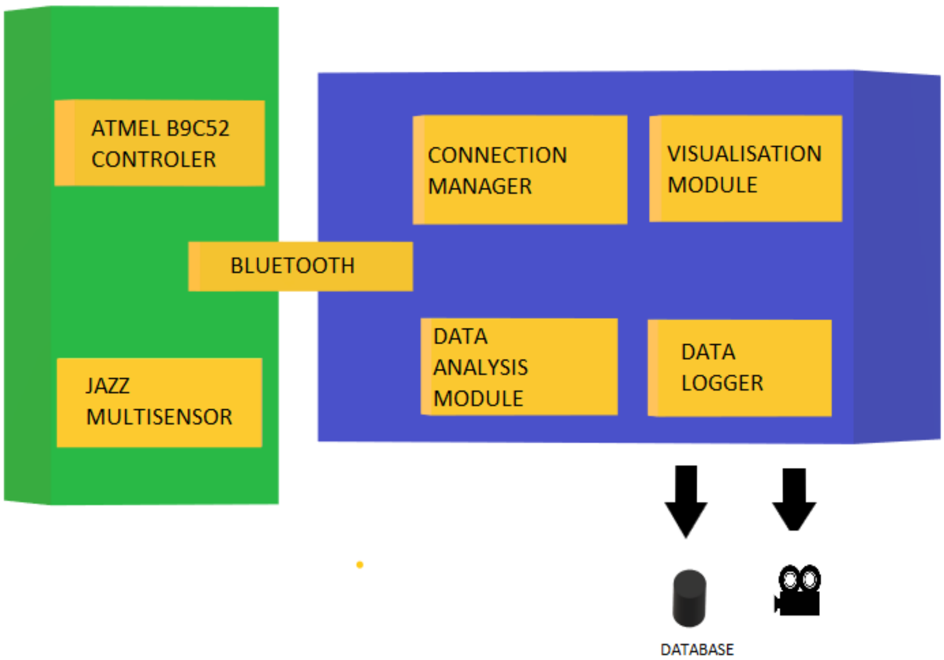


*Figure 1 The System Overview*

## 3.1. HARDWARE

The basic hardware components are discussed below:

### 3.1.1. Data Acquisition Unit

It collects physiological sensory data from the sensors via wireless communication over Bluetooth and is known as the mobile device for the system. The basic function is all about collecting data and sending it to Central Unit for further processing and analysis.

The RFID id cards are secured with a PIN which every operator need to use to get authorization of the DAU. The operator communicates with the DAU using a 5 key keyboard paired with an LCD smart display and a beeper. Whenever the system senses an abnormal situation an alert is fired on the display along with the beeps. The speech data is carried using a wired headset, connected to the DAU with a 3.5 mm usual audio port.

The Data Acquisition Unit consists several hardware components which are explained in detail along with the diagram.

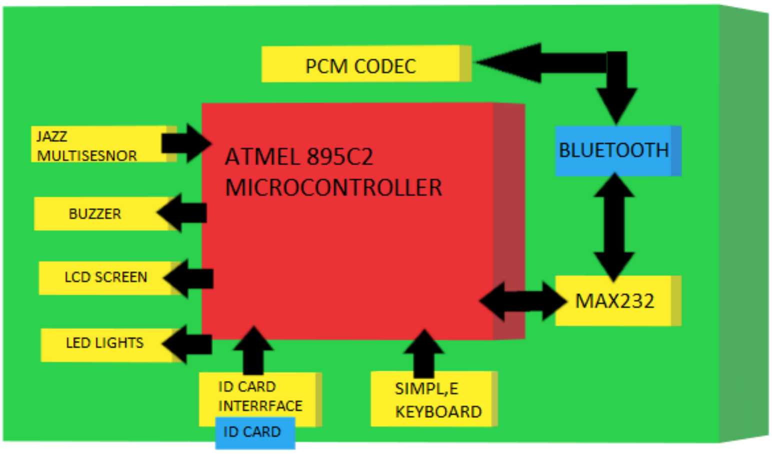


*Figure 2 Data Analysis Unit*

Hardware Specifications

1. The Atmel 8952 microcontroller is put to power and run the DAU due to its significantly low cost and high-speed serial port that makes data transfer and processing light blazing fast and so it is now has established itself as an industrial standard for blue eyes technology.

2. The Bluetooth module used is capable of synchronous voice data transmission which is also known as SCO link. The PCM codec is used to transfer operators voice and feedback from central unit. The PCM codec is considered to be very efficient in terms of data transmission over UART interface due to its tendency to lessen the amount of data and thereby reducing microcontrollers task. Additionally, Bluetooth module also performs voice compression in order to reduce the bandwidth used resulting in better overall sound quality.

3. The Jazz Multisensor is used as a sensor to track eye movements, ambient light intensity, blood oxidation level and acceleration along horizontal and vertical axis. The sensor is quite advanced as it sends 5.2kb of data every second with eye tracking data traced at 1kHz and other information at 250 Hz.

4. The alpha numeric LCD display provides the optimal place for giving information regarding events, PIN, wireless info and battery status alongside the LED. The keyboard is used to enter PIN and interact with the system physically.

Microcontroller Software Specification

The modelled software is based on a state diagram to facilitate easy debugging and testing and to keep the data small so as to maintain the highest program efficiency while using the lowest of the resource. This tedious task is handled by coding the DAU in 8051 assembler code. This in turn makes the implication easy. [3]

### 3.1.2. Central System Unit

It is basically the center for all data analysis and is also known as second component connected to wireless connection. The unit is equipped with Bluetooth on board and a codec written in PCM which is used for transmission of speech data. The module is associated to a PC via the USB connections in the series and parallel combinations. The speech data is attainable via usual 3.5 mm audio ports.

Connection Manager

Its basic operation is to manage wireless communications between the two modules. It performs tasks such as:

- Establishing communication with CSU
- Continuously look for devices in range
- Establish wireless Bluetooth relations
- Provide validated authorization to connections
- Buffer the incoming data
- To send the alerts to user

Data Analysis Module

This module is for the prelim’s scrutiny of raw data obtained from the sensors providing with information about the operator’s physiological condition. This module is made up of smaller analyzers registered at operator manager, extracting different types of information. Custom modules are created in order to fulfil and predict the supervisors need. This is done using a self-learning algorithm which checks for and verifies a data-set of previously recorded examples with recognition characteristic features.

The major thing is to detect and analyze saccadic signals and thus a saccade detector is used which recognizes eye movements which are generally very fast and collect accelerometer signal data for physiological constraints. The pulse rate scanner inscribes for ox hemoglobin and deoxy hemoglobin in the blood stream, these signals are sinusoidal in nature and are a direct indicator of heartbeat so, a pulse rate is determined on the basis of time delay between subsequent extremes of signal.

Data Logger Module

This module is designed to store the monitored data to allow the supervisor to collect, organize and scrutinize the operator's data. The module basis is to separate every user's data that is being recorded into an organized form.

Visualization Module

This module for a user interface wherein all the data from the operators is displayed in an initiative manner. The supervisor receives all the data in the form of pie charts which are representing operators total Visual time and active time on the system in percentages. It also compares any previously monitored data from the operator and display it alongside.

All the video and incoming sound source is available for the supervisor right in his hands along with the calculated operator's physiological condition. All messages and alarms are also let known to the user for better supervision.

## 3.2. SOFTWARE

Tools Used

A software which is called BlueDentist is used to check for connection establishments and management. The local devices are controlled and supervised using this piece of software for smooth operations.

Similarly, a simple dialog-based application for recording sensory information from Jazz Multisensor is used to read data via parallel ports and write program files to the operator's personalized identification cards using the standard parallel port and TTL compliance. [2]

# 4. EMOTION COMPUTING

Rosalind Picard in 1997 stated that it is quite important that the computers start to understand the emotions of humans for the computing community. Adaptive computing is the need for the modern times.

Affective computing has two important steps which are

- The ability to detect emotions
- The ability to express emotions

As Dryer and Horowitz stated that people with alike personalities acclaim each other collaborate well. People view and expect their computers to have a personality this way they will do well with their day to day jobs and interact well. Here’s why adaptive computing plays an important role.

## 4.1. THEORY

The researcher Paul Ekman who worked on facial expressions corelated emotional state of a person and his physiological parameters. He recorded participants measurements such as pulse, temperature, blood pressure, galvanic skin response (GSR), and somatic movement. He also managed to record the six basic emotions which were namely anger, fear, sadness, disgust, joy and surprise. The data was then analyzed using multidimensional scaling.

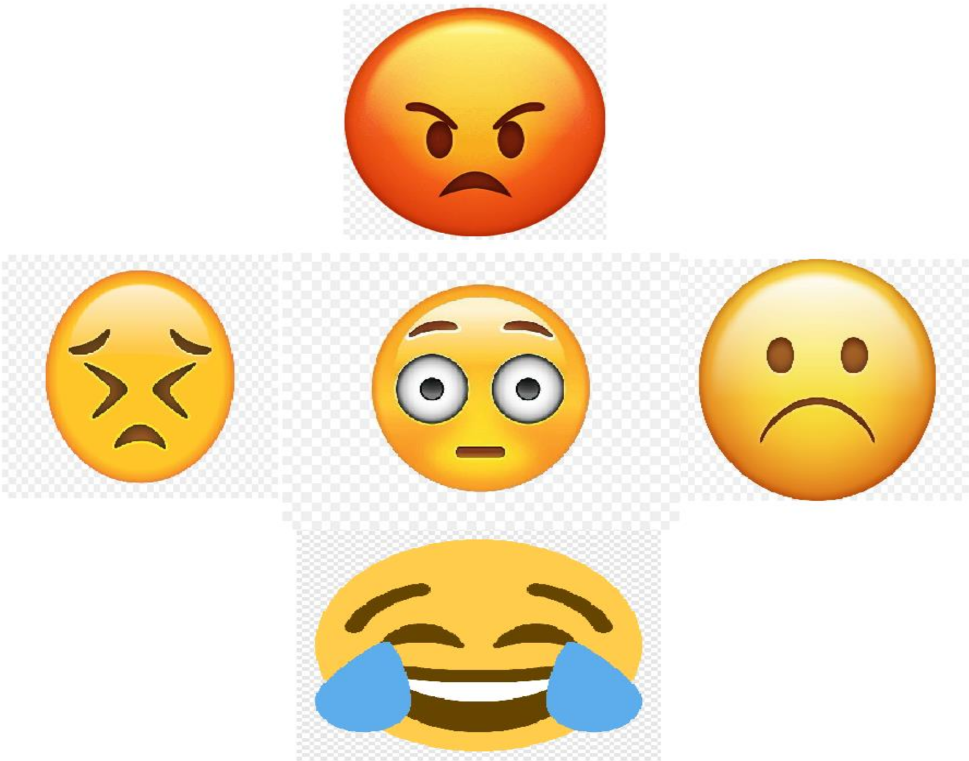

*Figure 3 The Human Emotions*

## 4.2. RESULT OF THEORY

The data is collected for each of the four physiological assessments which are namely GSA, GSR, pulse and skin temperatures and the six emotions which consist of anger, disgust, fear, happiness, sadness, and surprise and provided scores after an analysis. The scores thus analysed are compared with the baseline and difference among the two is found out and the scores which deviate more than one-and-a-half-point basis from mean are ruled out of consideration. A model is created using correlations among the physiological measurements and emotions which is then to be calibrated to the emotion mouse and the system. The signals received from the user are analysed and correlated and compared with the model. [3]

## 4.3. TYPES OF EMOTION SENSORS

Humans express their thoughts and emotions in different ways using their sensory abilities and so all the emotion sensors are designed around them only which are discussed below:

### 4.3.1. Emotion Mouse

The emotion mouse is the device used to obtain information through touch and act as an important input device. It can compute temperature, galvanic skin response, cardiovascular rates and involuntary movements and verifies them with six basic known emotional states. This helps in determination of users' mood. The mouse comprises of infrared sensors and temperature detection chips.

It can gather following information:

1) Behaviour
   a. Movement gestures
   b. Rate of mouse clicks
   c. Pressure applied on key press

2) Physiological information
   a. Cardiac rate
   b. User body temperature
   c. Electrical conductivity of the body

*Figure 4 Emotion Mouse*

### 4.3.2. Sentic Mouse

A redesigned handheld mouse which makes use of sensors to measure directional thrust for identification of emotional constraints such as likes and dislikes, attraction and avoidance.

### 4.3.3. Expression Glasses

It is a wearable device which is to be worn over the head covering the eyes this in turn allows for visualization of the interest of the user. The recent development includes learning automatically the interest area by continuous interaction with the computer. It can recognize favourite websites, bookmarks and open them according to time and day. [3]

#### 4.3.4. Eye Gaze

The eye gaze technology came into the existence in 1900s and have always been seen as the one of the vital parts in emotion computing and recent inventions and advancement provides assurance for a complete revamping in this idea of emotion-based computing.

# 5. FOUNDATION OF GAZE TRACKING

## 5.1. CLASSIFICATION OF EYE GAZE

Eye gaze studies and research about user's behaviour, intent, attention and cognitive process using several types of eye movements such as:

1. Fixations: It refers to the small periods in between the movements and visual inputs when the eyes are stationary. Parameters for measurement of fixation include variables such as total fixation duration, fixation spatial density, fixation rate, mean fixation durations, and fixation sequences.

2. Saccades: The rapid involuntary eye movements occurring between the fixations is known as saccades. The parameters for its measurements include fixation saccade ratio, saccade number and amplitude.

3. Scan path: The path followed by the eye gaze to reach a target point on a screen follows a chain of small infatuations that alternate in front of the operators' eyes is known as scan path. The parameters for its measurements include scan path direction, area covered, duration and length.

4. Gaze duration: The sum of all fixations made including the overall time of the user being consumed in each and every area over the areas of interest of the user is called gaze duration

5. Size of pupil: These are the parameters used to learn and measure cognitive nature of the workload. [4]

## 5.2. PRIMARY MODE OF GAZE ESTIMATION

These systems which are based on video fundamentally compromising of solo or multi camera setup, and well placed infra-red LED's. The most common and widely used gaze estimation is Pupil Centre Corneal Reflection aka PCCR method. The infra-red LED's are used to glimmer over the to get the cornea of the operators' eye. Images of the eye region and are usually working on 850 +/- 30 mm in order to achieve better contrast and colours. This in turn also avoids any unnecessary variations that could be induced by the natural light. Webcams that are used for video capturing are equipped with transmission filter to block any visible lights from outside environment. These cameras operate at 30/60 fps for better visual details.

There exist various types of UI for gaze tracing known as:

1. Active Tracking Interface
2. Passive Tracking Interface
3. Single modal Interface
4. Multi modal Interface

In active user interface, the operators eye movements are traced to initiate a computing function and information received by gaze tracking is utilized as an input for the model. Whereas the data collected is analysed for the user area of interest and attention with a non-commanding passive user interface. Single modal gaze makes use of only eye movements as the input interface while multimodal combines hardware such as mouse, keyboard and touch along with blink as its input interface.

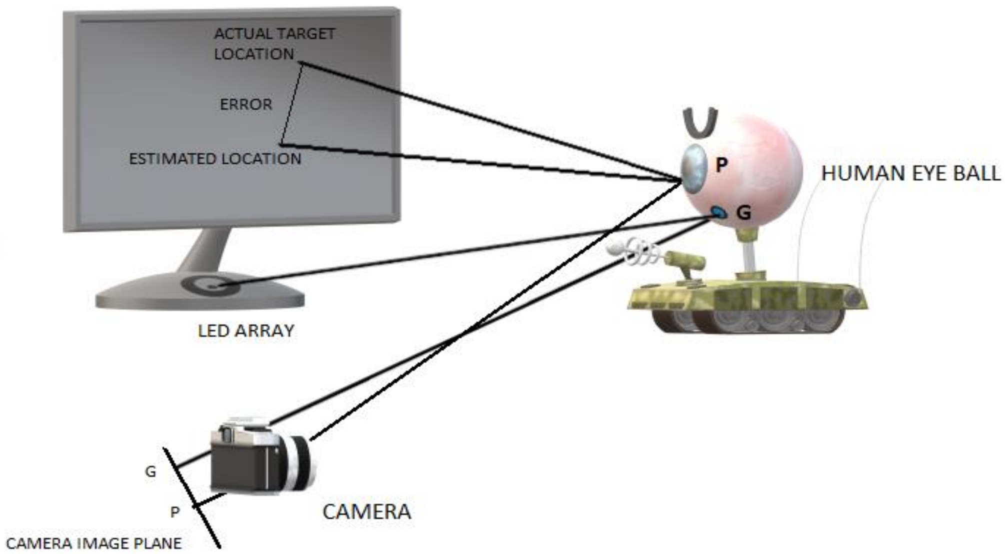


*Figure 5 Primary Setup for Gaze Estimation*

### 5.2.1. Calibration

Each person has different characteristics due to different gynematic syndrome and so the eye characteristics also varies from person to person. Therefore, the calibration becomes quite important. The internal vital parts such as retina and cornea are used for gaze estimation by determination of kappa angle. The process is carried out by web cameras which are capable enough to map the details on a data set when a user stares into them for some specified amount of time.

### 5.2.2. Eye gaze and Head position

Whenever a user changed his/her head orientation with respect to camera when he is still looking at some point on the screen the glimmer trajectory would change in order to the pupil centre of the user. So, the old calibrated data would become inaccurate to would give absurd results. This effect of change in head orientation has to be corrected and compensated before it gets further processed in gaze mapping algorithm. The algorithm that is used for this correction in head orientations is:

$$\mathbf{d_k - d_{kref} = k(d_{kgaze} - d_{kref})} \qquad \textbf{(6.1)}$$

where k refers to the parameter related to the head pan and tilt report
$d_{kref}$ = user gaze direction
$d_k$ = head pose direction
$d_{kgaze}$ = actual gaze direction

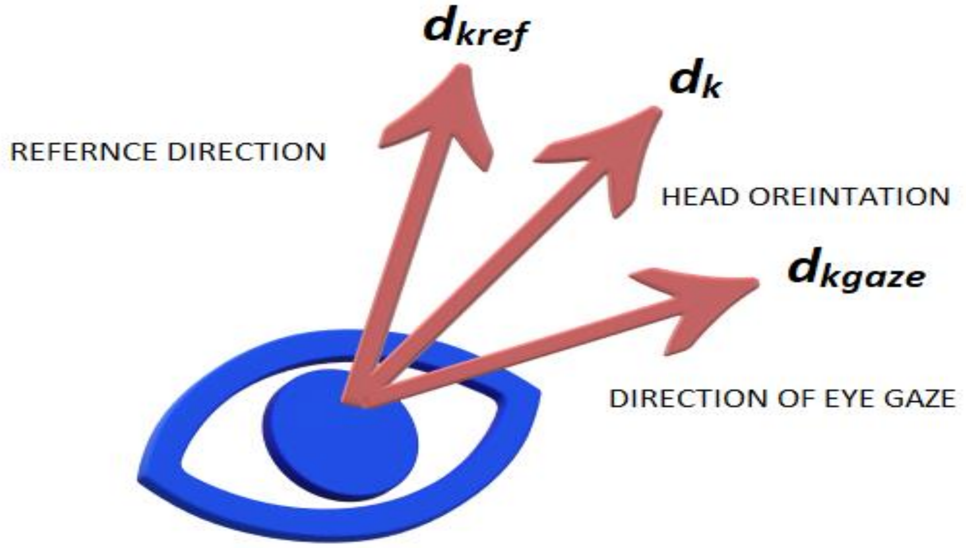


*Figure 6 Gaze Estimation and Calibration*

### 5.3. ESTIMATION OF EYE GAZE TRACKING

The typical eye gaze estimation techniques include an interface providing visual stimuli as the targets. The average variance between the measured gaze position of the system and the actual gaze position estimates the gaze tracking precision.

The eye gaze estimation is traditionally done for each one of the eyes but to avoid complexity similar computations are accounted for both the eyes. POG.$X_{left}$, POG.$Yl_{eft}$, POG.$X_{right}$, POG.$Y_{right}$ refers to the visual POG measured for each one of the eyes in X & Y coordinates. POG.X and POG.Y accounts for the separation of the eye from the main screen. Mean distance in the equation represents to the separation in between the eye and the tracker.[4]

Gaze point coordinates:

$$\mathbf{POG.X} = \frac{\mathbf{mean(POG.X_{left}+POG.X_{right})}}{\mathbf{2}} \quad \mathbf{(6.2)}$$

$$\mathbf{POG.Y} = \frac{\mathbf{mean(POG.Y_{left}+POG.Y_{right})}}{\mathbf{2}} \quad \mathbf{(6.3)}$$

Pixel accuracy (Pix_acc):

$$\mathbf{Pix\ acc} = \sqrt{}((\mathbf{target.X} - \mathbf{POG.X})^{\mathbf{2}} + (\boldsymbol{target.Y} - \boldsymbol{POG.Y})^{\mathbf{2}}) \quad \mathbf{(6.4)}$$

# 6. METHODOLY USED IN GAZE ESTIMATION

These estimation algorithms use NIR and corneal reflection to evaluate the gaze vector and use polynomial or geometrical functioning model to get the gaze estimation.

2D regression, 3D regression model and Cross Ratio based method are some of the several techniques falling into this category. There also exist some methods which use other statistics and details such as local features of the eye relating to shape and texture of the eye alongside the visible light to estimate the graze direction.

## 6.1. 2D REGRESSION BASED MODEL

This model map the gaze coordinates onto the main LCD frontal screen by use of estimation-based vector between the cornea and pupil centre by the means of polynomial transformation. The function used for mapping can be stated as:

$$\mathbf{f}: (\mathbf{x_e}, \mathbf{Y_e}) \to (\mathbf{X_s}, \mathbf{Y_s}) \qquad \textbf{(7.1)}$$

Where,
Xe, Ye & Xs, Ys: instrumentation and coordinates of the LCD screen respectively.

The relation is shown below:

$$\mathbf{X_s = a_0 + \sum_{p=1}^{n} * \sum_{i=0}^{p} a_{(i,p)} X_e^{p-i} Y_e^{i}} \qquad \textbf{(7.2)}$$

$$\mathbf{Y_s = b_0 + \sum_{p=1}^{n} * \sum_{i=0}^{p} b_{(i,p)} X_e^{p-i} Y_e^{i}} \qquad \textbf{(7.3)}$$

Where,
n: order of the polynomial
$a_i$ and $b_i$: coefficients

The user is then asked to gaze at different location on the frontal screen for calibration and polynomial optimization. So as to reduce the mean squared difference (ε) of the actual physical LCD screen coordinates and the computed coordinates the order of the equation and coefficients are chosen which is given as:

$$\boldsymbol{\varepsilon = (X_s - M_a)^T (X_s - M_a) + (Y_s - M_s)^T (Y_s - M_b)} \qquad \textbf{(7.4)}$$

Where,
a and b: vector coefficients
M: matrix of transformation

$$\boldsymbol{a^T = [a_0 a_1 \ldots a_m], b^T = [b_0 b_1 \ldots b_m]}$$

$$M = \begin{bmatrix} 1 & X_{e1} & Y_{e1} & \cdots & X_{e1}^{n} & \cdots & X_{e1}^{n-i}Y_{e1}^{i} & \cdots & Y_{e1}^{n} \\ 1 & X_{e2} & Y_{e2} & \cdots & X_{e2}^{n} & \cdots & X_{e2}^{n-i}Y_{e2}^{i} & \cdots & Y_{e2}^{n} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \cdots & \cdots & \cdots \\ 1 & X_{eL} & Y_{eL} & \cdots & X_{eL}^{n} & \cdots & X_{eL}^{n-i}Y_{eL}^{i} & \cdots & Y_{eL}^{n} \end{bmatrix} \quad \textbf{(7.5)}$$

m: total number of the coefficients
L: Total calibration points in number

When the matrix M is inverted the coefficients of the equation are obtained with equation

$$A = M^{-1}X_s\,, b = M^{-1}Y_s \quad \textbf{(7.6)}$$

A new approach is now been taken into account which includes only 2 IR LED in order to make the algorithm more suitable. The POG is estimated via mapping functions and calibration process.

**6.2. 3D REGRESSION MODEL**

This method makes use of the internal features of the eye which are called geometrical features for the estimation of the operators' gaze. It can be classified on the basis of number of camera setup being used such as a solo or multi and the methodology being used for the camera calibration. Single camera systems are quite simple and are faster in re-acquisition capabilities.

Multi camera methods can go a step further and achieve quite high accuracy for head movements but due to the complex procedures will require an elaborate calibration of the system with different LED positions and camera for 3D measurements.

**6.3. CROSS RATIO BASED METHODS**

This method is well known for its 3D modeling by the use of LED lights and their reflections produced over the eye lens. It makes use of NIR lights which are supposed to be situated on the corners of the LCD screen. The estimation and computing task for gaze tracking is done through projections produced by the eye lens and cornea which are projected over to the camera's imaging plane. The distance and data are thus analyzed for gaze tracking. [4]

**6.4. APPEARANCE BASED METHOD**

In this method basically the eye region data is analyzed for feature extraction from eye images. The shapes and texture variations are represented and trained by a statistical model.

The shape vector is stated as:

$$s = (x_1, x_2, \dots, x_L, y_1, \dots, y_L)^T \quad \textbf{(7.7)}$$

Where,
L: total number of points of landmark

The Principal Component Analysis is conducted over the aligned shape data set to obtain the shape model as:

$$s = \bar{S} + \varphi_s b_s \tag{7.8}$$

$$\bar{S} = \frac{1}{N_s} \sum_{i=1}^{N_s} s_i \tag{7.9}$$

Where,
$\bar{s}$: mean shape vector
$N_s$: shape observation number
$\varphi_s$: eigen vector matrix
$b_s$: shape parameters set

Therefore, following the above criteria for each training texture vector is given out as:

$$t = (t_1, t_2, \dots, t_p)^T \tag{7.10}$$

P: total number of texture samples.

Over the texture vector, a texture model is obtained through the PCA which is given as:

$$t = \bar{T} + \varphi_t b_t \tag{7.11}$$

$$\bar{T} = \frac{1}{N} \sum_{i=1}^{N_t} t_i \tag{7.12}$$

Where,
$N_t$: number of texture observations
T: mean texture vector

The shape set along with the texture parameters ($b_t$) express the appearance ability of the model:

$$c = \begin{pmatrix} w_s b_s \\ b_t \end{pmatrix} \tag{7.13}$$

The algorithm known as Active Appearance Model is considered most suitable to create this statistical model of a new eye image.

The models are tested, trained and mastered to for whole face and eye regions for both the global and local appearance.

## 6.5. SHAPE BASED METHOD

This method makes use of deformable eye templates which are formed using circles of the iris, two eye contours parabolas which are then fitted to an eye image. This way the similarity is found out between the template and image of a region. Mean square method, cross correlation method and a modified version of it are used for calculations. The intensities of the pixel of template is shown as T(u,v) and the captured image pixel intensity is represented by I(I,j) and the resemblance between image and template is given as S(I,j). For the cross-correlation method S is given as:

$$\boldsymbol{S(i,j)} = \frac{<T\ X\ I_T> - <T><I_T>}{\sigma(T)\sigma(I_T)} \tag{7.14}$$

Where,
< >: average operator
< x >: pixel-by -pixel product

$$< \boldsymbol{T} > = \frac{1}{n}\sum_{u,v} \boldsymbol{T(u,v)} \tag{7.15}$$

$$< \boldsymbol{T\ X\ I_T} > = \frac{1}{n}\sum_{u,v} \boldsymbol{T(u,v)I(i+u,j+v)} \tag{7.16}$$

σ: standard deviation of the region that is being resembling

$$\boldsymbol{\sigma^2(T)} = \frac{1}{n-1}\sum_{u,v}\left(\mathbf{T(u,v)}\right)^2 - < \mathbf{T} >^2 \tag{7.17}$$

The mean squared error in the methods similarity measure is given by:

$$\boldsymbol{S(i,j)} = \frac{1}{n}\sum_{u,v}\left(\boldsymbol{T(u,v) - I(i+u,j+v)}\right)^2 \tag{7.18}$$

# 7. MAGIC POINTING

No matter how much the eye tracking technology matures there will be imperfections to existing gaze tracing techniques and the two main observed such outcomings are:

1. Eye gaze is not suitable enough to manage UI widgets due to the precision issues, scroll bars, hyperlinks and slider bars on the modern GUI due to the tiny size of the fovea and unexpected subconscious involuntary motions produced by the eye. A .044 in of arc is produced for one degree which is almost double the size of scroll bar and not even comparable to size of a typical character for viewing distance of 25-inch.

2. The eye movements are sometimes voluntary whereas at other times maybe dependent on the external events as eye being the primary visual organ has not evolved to be develop into a command organ. One has to be conscious about his/her gaze and the interval for the looking at the one object and the eye blinking. The eye gaze requires a minimum stare of 200ms to select an object but if users fails to do so the object would not be selected and if he stares for more than that the object would get selected regardless of choice. This can turn annoying and time consuming for most users. [3]

MAGIC pointing is an abbreviation used for "Manual and Gaze Input Cascaded pointing" which dynamically combines eye gaze technology with manual inputs to acquire the desired results with more accurately. The bottom-line idea is to apply gaze to redefine the location of the pointing cursor to a home position in the vicinity of the target to which the user is gazing at. This will reduce the manual efforts and pointer movement. The liberal and the conservative are the two major pointing techniques for target spotting and pointer placement used in the MAGIC pointing

The liberal approach talks about to wrap the cursor around the object where the user gaze is set up and then the user can take up a manual input to select the target object. The cursor wraps around at a distance of around 120 pixels which is sufficient distance for easy selection by manual input. The approach is quite proactive since cursor waits readily in the vicinity of the target the system sees as potential target. Though the user may find this a bit of too sensitive but will adapt to it easily.

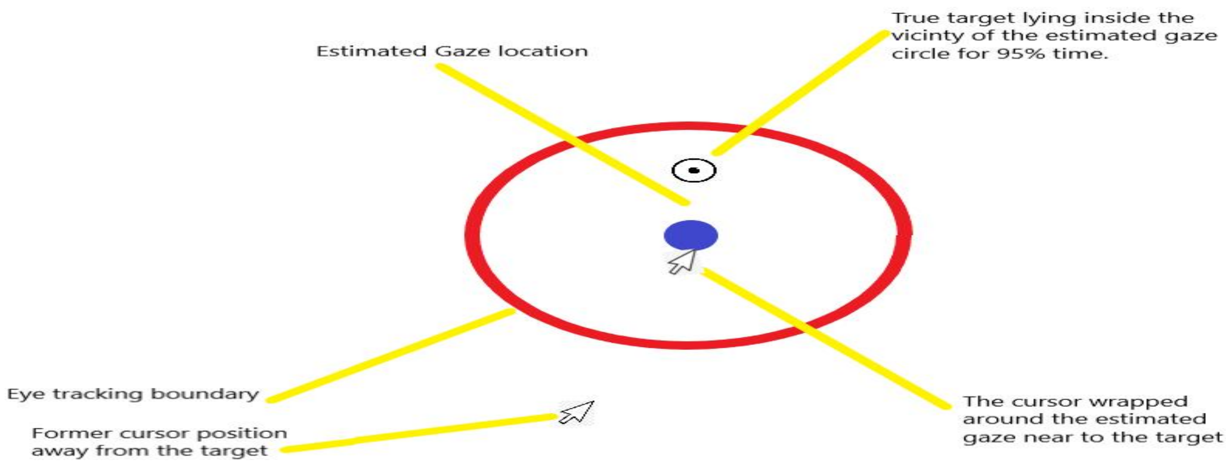

*Figure 7 The Liberal Approach for MAGIC pointing*

The conservative approach is somewhat different as it does not wrap the pointer around the target until a manual command input is acquired by the system. Once the system gets a go by manual the pointer is wrapped and placed near to the target area. The goal is all about getting the pointer out of the blue when a manual input is given.

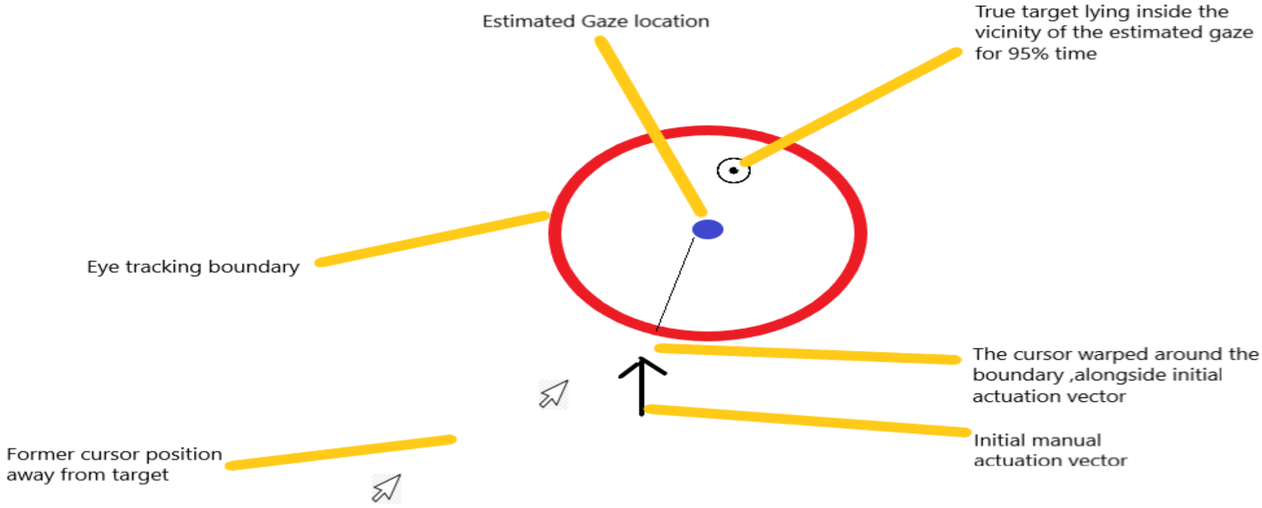


*Figure 8 The Conservative Approach for MAGIC pointing*

The potential advantage of liberal and conservative pointing techniques are:

1. As the cross screen long distance cursor gets out of the picture of manual control the manual stress and fatigue gets reduced.

2. The traditional eye gazed linked along with manual inputs completes the task the accuracy levels increase drastically.

3. It works more naturally and the user does not always need to be aware of his gaze because it will always require some manual inputs.

4. The pointing becomes faster than manual pointing as it uses eye gaze technology to do it.

5. As the pointing are by manual efforts reduce the user may feel that system is working fast and more efficiently than ever and that gives pleasure to the user.

The fourth point discussed above is the result of Fitts' Law which states that "manual pointing time is logarithmically proportional to the A/W ratio, where A is the movement distance and W is the target size." In layman's language the target which will be too close or too far will take up longer period of time to acquire. Since the target area stays the same as before for MAGIC pointing the distance of cursor movement is reasonably shortened. [3]

There might be some issues that would relate to magic pointing such as delays, errors and inconvenience such as:

1. The cursor for the liberal approach can be seen as overactive sometimes and might

become quite distracting for the user, especially when he is trying to read. As the user's eye will move further than pre-set gaze of 120 pixels the pointer would move with it. Though this can be cured by introducing new constraints so that the system knows when a user is reading.

2. For the conservative approach there might exist some uncertainty about pointers location at which cursor may appear around a target. This will create problems for novice to adapt to it and to use more of the manual method which will prolong the action time. Sometimes the user may find himself waiting for cursor to appear before he can start using it. So, the user will need to develop novel hand eye coordination.

3. With the manual inputs in hands the user always knows location of his pointer and therefore can perform parallel motor action but with the eye gaze in use there will be a hindrance in parallel action process.

# 8. SUITOR

SUITOR stand for Simple User Interest Tracker. This a technology developed by the IBM to grow upon the idea of Blue Eyes System. Computers have grown much more powerful and rich in terms of performance and with the perceptual abilities and sensory abilities can recognize different human emotions and interest. The use of emotion and septic mouse and eye glass has enabled to read human brain and provide a intimate relation between the computers and the user. And the SUITOR provides with a revolutionary approach in this movement.

The SUITOR usually tracks user inputs and eye gaze and may help extract and display more of the information to the user. For example, whenever a user is browsing a webpage the SUITOR deeply observes and track the interest areas of user. This creates an intimate bond between the user and computer. The suitor can take information through a handheld device and respond to it back. It can keep track of applications and web pages running and adapt according to user to help him out best in his work. SUITOR focuses on exploiting the non-verbal ways of communication with the user. [3]

# 9. SPEECH RECOGINITION USING ARTFICIAL INTELLIGENCE

Human Beings communicate using both verbal and nonverbal means but the verbal communication means simply is still the best way to understand the human emotions as they provide you with the most accurate psychological insights of the user. An efficient speech recognition system would only make the interaction between computer and human more intimate and friendly. The variations in various prosodic parameters of a natural language can help identify emotions. The list of useful features of a speech include voice quality, speech rate, timing, utterance intensity and contour.

Algorithms based on two basic approaches namely: Pattern Recognition Approach and Artificial Intelligence Approach. Algorithms which are based on Pattern Recognition are Template matching and that on artificial intelligence include knowledge sources, stochastic of speech signals. The important approach is based on Stochastic modelling using hidden markov models. Among this the most popular algorithm is the template-based designing. The speech recognition process has front end processing part which converts speech signals into a usable form. The first step being feature extraction at the very beginning since it has features that are literally of no use to the talker and channel variability. The features obtained thus decrease the data rate for further processing. [5]

The speech recognition process is quite likely to affected highly by noise present in the system surroundings. Noise is the major source of obstacles for the system which hinders speech recognition. Feature extraction plays an important role in the flawlessness of the system. A system may contain minimum of three filters and as the number of filters goes higher the probability for the precise recognition goes up.

## 9.1. HIDDEN MARKOV MODEL

The hidden markov model is a stochastic capable after it learns a phase to estimate likelihood of observation sequence. It can be defined by all the following parameters:

N: the number of model states
A = $\{a_{ij}\}$ = P(qj/qi): is a matrix of size N*N

It is characteristic transition matrix between states of model. The probability of transition of state j depends on state i

$$\boldsymbol{p\left(q^t = \frac{j}{q_{t-1}} = i, qt-2 = k, \ldots\right) = p(qt = \frac{j}{qt-1} = i)} \qquad \textbf{(10.1)}$$

$$\boldsymbol{B = \{b_j(o_t)\} = P\left(\frac{o_t}{q_j}\right)} \qquad \textbf{(10.2)}$$

Where j [1,N]

It is the set of emission probabilities of the observation $o_t$ when system is in state $q_j$. A continuous probability is defined by the below relation

$$b(o,m,v) = N(o,m,v) = \frac{1}{\sqrt{(2\pi)^n|c|e^{(-1/2(o_n-m_i)c^{-1}(o_n-m_i)}}} \qquad \textbf{(10.3)}$$

Where
o: observation frame
C: covariance matrix

$$C = \frac{1}{n} - 1 \sum_{k=0}^{n}(o_k - m_k)' * (o_k - m_k)c \qquad \textbf{(10.4)}$$

m: the mean of each coefficient

$$m = \frac{1}{n}\sum_{k=1}^{n} o_k \qquad \textbf{(10.5)}$$

Multi gaussian probability density is used for different pronunciation of words given as:

$$B_j(o_t) = \sum_{i=1}^{k} C_{ij} * b_j(o_t) \qquad \textbf{(10.6)}$$

K: number of gaussian
$C_i$: gaussian weight of I in j

**9.2. DYNAMIC PROGRAMMING**

Dynamic time wrapping algorithm is suitably used in speech recognition as it measures resemblance between two chain or sequences may vary with time and speed. The principle is to compare two speech signals based on distance between two matrices corresponding to coefficients of mcl of the two signals.

Euclidean distance between two signals is measured using

$$d(i,j) = \sqrt{\sum_{k=1}^{k}(x_k - y_k)^2} \qquad \textbf{(10.7)}$$

Where, i = $\begin{pmatrix} X1 \\ \vdots \\ XN \end{pmatrix}, j = \begin{pmatrix} Y1 \\ \vdots \\ YN \end{pmatrix}$

Then calculate a minimum distance by traversing the element of the matrix obtained using relation:

$$g(i,j) = \min \begin{cases} g(1-1,j) + d(i,j) \\ g(1-1,j-1) + 2.d(i,j) \\ g(i,j-1) + d(i,j) \end{cases} \qquad \textbf{(10.8)}$$

The final distance is:

$$\boldsymbol{G} = \frac{\boldsymbol{g(I,J)}}{\boldsymbol{I}} + \boldsymbol{J} \qquad \textbf{(10.9)}$$

Where: I,j: length acoustic arrays corresponding two signals.

# 10.APPLICATIONS

1. It can be used in various fields including defense and security where facial recognition can be used to recognize criminals and offenders. The system can also be used to judge one's physiological condition and help curb crime. The weapon operators in defense forces can use speech recognition to fire weapons and focus their attention solely on control and target.

2. The system can be used in cars to understand the driver's physiological condition and advise him on the basis of his emotional state on his driving style or stop him from driving when he isn't in a state to drive. It can also trigger alarm when the driver feels dizzy or doze off while driving.

3. Aircrafts and pilots can implement this to make the process less complicated and manage aviation easily.

4. Offices and work places can use this system to gather information about their staffs' emotional state and physiological condition so that they can provide necessary help and focus them to work efficiently.

5. The computer can be made an intimate partner for a user which will motivate them to work better without any distractions and the user can perform multiple tasks without actually putting any manual efforts to it.

6. The business can use the information to display advertisements and web pages according to user interest.

7. Many smart devices such as Mobiles, Television and home appliances are using more of this technology to enhance their product experience.

# 11.CONCLUSION

The blue eyes system will make our lives much easier and more fun. You would see people interacting with their computers just like normal human beings. People who feel they are lonely and fail to find someone to listen to them would now have a companion. This would also help curb suicides, anxiety attacks and emotional displeasure. Moreover, the system will help reduce accidents, fatigues, muscle loss and other serious problems. The human world will have a new friend which maybe personal for everyone who would understand them. The only disadvantage being that it would not be able to make people change their thoughts.

# 12.REFERENCES

[1]. Shehu, Nura Muhammad, “The Concept of Blue Eyes Technology”, International Journal for Research and Development in Technology, Vol-4, Issue-5, pp 21-22,2015.

[2] Upasna Arya, “Human Brain Interaction and Motion Perception using Blue Eyes Technology”, International Journal on Emerging Technologies, Special Issue NCETST, pp 454-456, 2017.

[3] Binyamin M, “Blue Eyes Technology”, Division of Computer Engineering School of Engineering Cochin University of Science & Technology, pp 2-13, 34, 2010.

[4] Anuradha Kar, “A Review and Analysis of Eye-Gaze Estimation Systems, Algorithms and Performance Evaluation Methods in Consumer Platforms”, IEEE Access, pp 1-8, 2017.

[5] V.V. Nanavare, S.K. Jagtap, “Recognition of Human Emotions from Speech Processing”, Procedia Computer Science, Elsevier B.V., pp 28-30, 2015