IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-12, Dec 2019 E-ISSN: 2347-2693

Detection of Facial Micro Expressions and Textual-Tracking for Paralyzed Using Computer Vision

Sameer G Mathad[1*], Pruthvi L R[2], Sanjana C S[3], B Akash[4], Deepthi V S[5]

[1,2,3,4]Dept. of Information Science and Engineering, Global Academy Of Technology, Bangalore, India
[5]Visvesvaraya Technological University, Belagavi, India

[*]*Corresponding Author: sameergmathad@gtmail.com Tel.: 9972899815*

DOI: https://doi.org/10.26438/ijcse/v7i12.6266 | Available online at: www.ijcseonline.org

Accepted: 13/Dec/2019, Published: 31/Dec/2019

***Abstract*—** Paralysis is the loss of voluntary muscle control. Many people cannot move a single part of their body. Even though the paralyzed people are cognitively aware of their surroundings, they have no means of communication. The paralyzed people have lost their ability to talk, type, etc. These victims affected by Locked-In Syndrome have their thoughts and ideas trapped inside of them. Usually people with paralysis have total control over their eye movement. Therefore this project aims to build a real time interactive eye blink system that allows paralyzed people to easily express themselves. Some people who are affected with Amyotrophic Lateral Sclerosis (ALS) and Tetraplegia can also express their emotions (Eg: smile) and hence these expressions can be used as commands for performing some helpful tasks. We aim to show this by implementing sentiment-based music system to play music by retrieving their current mood via video streams.

***Keywords*—** Formatting real time interactive eye blink system, Locked-In Syndrome, Amyotrophic Lateral Sclerosis (ALS), Tetraplegia, sentiment-based music system and video streams.

I. INTRODUCTION

Paralysis is a disease which is dramatically more widespread all over the world. There are about 1.7 percent of the people in total U.S. population, or about 5,357,970 people reported they were living with some type of paralysis, which is defined in the study as a central nervous system disorder results in difficulty or inability to move the upper or lower extremities [1].
The leading cause of paralysis was by stroke (33.7 percent), another may include a spinal cord injury (27.3 percent) or they may be affected with multiple sclerosis with a percentage of 18.6. According to the global study, there are about 1 out of 50 people living with paralysis – which sums upto almost 5.4 million people (Figure 1).

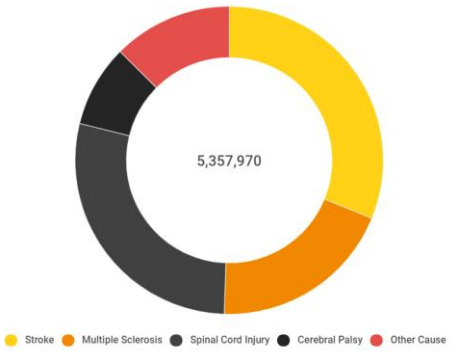

Figure 1. Statistics for cause of Paralysis

Locked-in syndrome is one of the types of paralysis which occurs when a person is quadriplegic and also has no way to produce speech or facial movements. The patients affected will not be able to communicate but is aware of their surroundings. The patient can hear and see and has normal intelligence and reasoning ability but can communicate only through eye movements.

Locked-in syndrome is mainly caused due to brainstem hemorrhage or infarct; rarely other causes such as trauma, tumors, or infection, for example, may cause the syndrome.
Locked-in syndrome has many signs and symptoms which include quadriplegia and the inability to speak or make facial movements, but those affected are aware and may communicate by eye movements [2].

Tetraplegia refers to a spinal cord injury above the first thoracic vertebra, or which is present within the cervical sections of C1-C8. The result is some degree of paralysis in all four limbs—the legs and arms. The degree of paralysis varies depending on the nature of the injury, the extent to which you've undergone rehabilitative therapy, and oftentimes on factors that are not yet well-understood—maybe even a bit of luck.

There are approximately about 12,500 new spinal cord injuries each year, meaning 40 out of every one million people suffer such an injury (Figure 2).

Between 240,000 and 337,000 Americans currently live with a spinal cord injury [3].

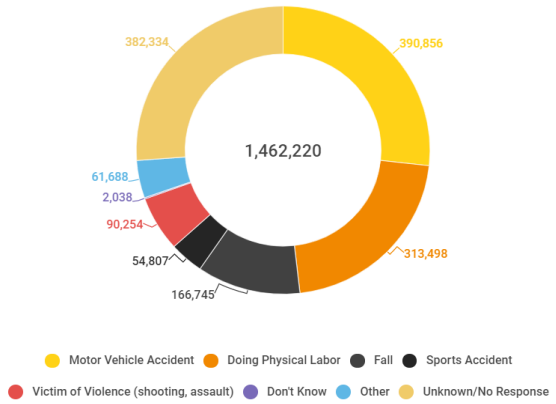

Figure 2. Statistics for cause of spinal cord injury

II. RELATED WORK

Few existing systems are already deployed in the public environment, but is not been developed for helping paralyzed people separately. One of the existing system works by comparing the captured eye image with already present images in the database. As the video frames are captured continuously comparison process becomes very slow and performance decreases. This process is not for communication and hence doesn't help paralyzed people to convey their messages. So our goal is to build an interface that captures live images and applies TSL color space to differentiate the blinks for executing the commands. And since the blinks are identified, we aim to use this method for creating our interface for Blink-To-Text communication, where a word is suggested for a character or the word can be formed explicitly and displayed on the screen that corresponds to the blink. This method can be used for paralyzed people as they have lost the ability to speak and use their eye blinks to convey their thoughts, i.e. as a mode of communication. And through this the care takers can easily understand their thoughts not only when they are with the patient but also away as the message formed can be sent as a text or an email.

Some existing music systems using face detection extracts the mood of the person by capturing the picture and plays a music playlist which is already classified and categorized for a particular mood which is played over and over again. Here the songs are classified and not the lyrics which may be insufficient for categorization process, decreasing the overall performance. In our Implementation we try to generate a music playlist by classifying the lyrics of the song based on the mood extracted from the user and play or provide the link of the songs. The new set of songs i.e. the playlist is suggested every time the mood is extracted.

The main objective is to implement computer vision and algorithms to help a patient affected with either of the following diseases: Tetraplegia, Amyotrophic lateral sclerosis (ALS), Locked- In Syndrome through our software. For the simplicity the proposed system can divided into two Interfaces:

1) Blink-To-Text Communication:
Step 1: Image/video acquisition from the camera.
Step 2: Convert video to frames.
Step 3: Classifying eye blinks for retrieving particular interface.
Step 4: Process following blinks for corresponding characters to suggest appropriate words.
Step 5: Display and Send the formed message either through an email or local message.

2) Sentiment Based Music System:
Step 1: Image/video acquisition from the camera.
Step 2: Convert video to frames.
Step 3: Classifying eye blinks for retrieving particular.
Step 4: Identify and Extract the emotion of the person and allocate to a particular mood category.
Step 5: Based on this mood a music playlist is generated in the database that is played.

Ramya Ramanathanet.al [4] proposed an intelligent agent that sorts a music collection based on the emotions conveyed by each song, and then suggests an appropriate playlist to the user based on person`s current mood. The user takes a picture of themselves at that instant. This image is subjected to facial detection and emotion recognition techniques, recognizing the emotion of the user, to recommend the playlist.Retrieves music that has been stored on a system, and processes this music by using various Python packages. Once musical feature extraction has been completed, the music collection can be clustered using any efficient algorithm. Then classification is performed.The emotion detected from the image processing is given as input to the clusters of music to select a specific cluster.

Atish Udayashankaret.al [5] describes in their paper that the main aim is to design a real time interactive system that can assist the paralyzed to control appliances, playing pre-recorded audio messages, through a predefined number of eye blinks. Face tracking can be done by using a set of trained Haar cascade classifier, and a template matching technique is employed to track the eye. In Contour extraction, a set of 16 landmarks are created at regular intervals. Eight points are used to represent each eye. Then the Gabor filter is used to extract arcs of the eye. The image of the eye is first threshold and later a median blur filter is applied to it. The resultant image obtained after applying the filtering shows a clear difference between the open and closed eye, and hence helps in identifying eye blinks.

Swathi Chauhanet.al [6] proposes that the work is to analyse the lyrics of Hindi-language based songs, in order to detect the mood of the listener. The songs are reduced to such a level so that only the relevant words are used for mood detection. By using mood taxonomy, songs can be distinguished between happy or sad. Data is applied to Latent Dirichlet Allocation model to discover the hidden emotions within each song. Topic model is a statistical model used for identifying clusters in a corpus of data. It is basically a method for tracking the cluster of words in large bodies of text. A song can display varying emotions. During collection of the corpus, it was seen that most of the songs display positive emotion such as excited, pleased, relaxed etc and also negative emotions such as afraid, frustrated, depressed etc. For simplicity, we divided this taxonomy into two classes.

Leo Paulyet.al [7] presents a novel method for eye tracking and blink detection in the video frames obtained from low resolution consumer grade web cameras. The method presented is non-intrusive and hence provides a comfortable user interaction. The accuracy of the method eye tracking is about 92.3% and that of blink detection method has 92.5% when tested using standard databases and a combined accuracy of 86% which was tested under real world conditions of a normal room. The method used involves Haar based cascade classifier for eye tracking and a combination of HOG features with SVM classifier for eye blink detection. The process includes the steps Capturing video frames, Face Detection and Extraction, Eye Region Extraction, Eye detection, Feature Extraction and Classification.

Beatriz Remeseiroet.al [8] presents a methodology to perform the automatic detection of eye blink using consumer videos acquired with low-cost web cameras. This method includes the detection of the face and eyes of the recorded person, and then it analyses the low level features of the eye region to create a quantitative vector. Finally, vector obtained is classified into one of the two categories considered open and closed eyes by using machine learning algorithms. Since the proportion of skin in a closed eye is greater than that of open eye, the TSL color space has been considered due to its effectiveness in skin analysis. Supervised learning involves learning a mapping between a set of input features and output labels. The resulting classifier is then used to assign class labels to the new instances and Naive Bayes, Random tree, Random forest, Support vector machine, Multilayer perceptron algorithms are used.

Bharadwaj Raghavan [9] describes in his paper that the developers introduced “burst capture”, the ability for your camera to take photos in rapid succession over a short period of time, allowing users to select the best photo from the set. Unfortunately, the vast majority of burst capture photos are unnecessary, taking up large swaths of memory on the mobile device. It implements an adaptive template based blink detection algorithm which detects frames where people are blinking so that they can automatically be deleted from your device. The initial implementation has an F1score of 94.39% and is capable of running in real time on a mobile device. The program usually has one or many templates of open eyes gathered and it can determine whether a person is blinking by computing a similarity score. Based on this, the program classifies the image as “blink” or “non-blink”. Open eye templates are created in real time, to have more personalized templates. They use frame differencing and binary thresholding while a person involuntarily blinks to locate the eye regions and grab templates after the blink occurs.

Tereza Soukupova´et.al [10] describes in their paper that a real-time algorithm to detect eye blinks in a video sequence from a standard camera is proposed. The landmark detectors which are trained using wild datasets show excellent robustness against a head orientation with respect to a camera, varying illumination and facial expressions. Thus the algorithm proposed determines the landmark positions, extracts a single scalar quantity – eye aspect ratio (EAR) – and characterizing the eye opening in each frame. Finally, an SVM classifier is used to detect the eye blinks as a pattern of EAR values in a short temporal window. The simple algorithm exceeds the state-of-the-art results on two standard datasets. The proposed SVM method that uses a temporal window of the eye aspect ratio (EAR), outperforms the EAR thresholding. The SVM algorithm runs in real-time, since the additional computational costs for the eye blink detection are negligible besides the real-time landmark detectors.

Dogukan Aksu et.al [11] proposes a real time vision system is presented to provide a communication way to people who has severe disabilities. Patients can choose words on an alphabet tree which is designed on a binary tree by blinking right and left eye, thus they will make sentences. The template matching algorithm is successful for all conditions such as detecting face and eyes together, detecting image in enough light with and without glasses, detecting image in a bad lighting condition with and without glasses. The template matching algorithm is used to decide whether the eyes are open or not by comparing the eye images which are on template with current eye images. If the similarity of two patterns comes below of a threshold value, it can be understood that there is a blink.

Kruti Goyal et.al [12] describes in their paper that the intention is deep study of face detection using open CV. A tabular comparison is performed in order to understand the algorithms in an easier manner. It talks about various algorithms like Adaboost, Haar cascades. This paper aims to help in understanding the best prerequisites for face detection.The speed of openCV uses c/c++ libraries which directly provides the computer with the machine language code and helps in faster execution whereas on MATLAB, the

computer initializes by interpreting the code and converting it into java and finally executes the script. The portability and cost of openCV is better than matlab. When compared with algorithms of haar cascade, camshift and finding via motion it was found that haar cascade algorithm was best suitable for identifying the face region.

Veena N et.al [13] describes in their paper that it is nearly impossible for such paralyzed people to communicate with others. This system allows physically handicapped patients to convert their intent to text using their eye-blink in order to use a computing system motor capability. Paralyzed people can use their eyes to communicate through eye blink. This eye-blink can be used to communicate information. After detecting the face, the next phase is the localization of eye region in the input. Location of the eyes is identified from the geometrical requirements. The conventional method tells that the eyes will be situated at 0.4 away from head top. If the time period of eye-blink detection is more than 250 ms and less than 2 s, then such a blink is said to be voluntary one.

Mu-Chun Suet.al [14] proposes an implementation of a low-cost vision-based computer interface which allows people with severe disabilities to use eye blinks to access computers and communicate with other persons. This communication aid requires only one low-cost web camera and a personal computer. Many experiments were conducted to test the performance of the proposed eye-blink-based communication aid. Frame differencing technique is the method used for eye blink detection. For current image, frames are used as the first four templates to determine the possible position of the eye region in the next image frame. Simple approach for generating the template for the eye region is to use a box around the centre of the working eye region. These four templates are used for eye tracking.

Yajie Huet.al [15] describes in their paper, a new framework to identify accurate social tags of songs to clean and filter music tags. Then apply an improved hierarchical clustering algorithm to group the tags to build a tag category. Based on the category, classify music songs using lyrics. In order to know the semantic information of lyrics, apply CLOPE to cluster lyrics and use the centroid of the corresponding cluster to represent the lyrics. Based on the Naive Bayes method, the probability of assigning lyrics to particular class is predicted. The classification result is then used for determining whether a social tag is accurate. The experimental results exhibits that the proposed framework is effective and encouraging. The category is automatically generated from tags. Compared to a pre-defined category, an auto-generated category is better in representing the songs be-cause it mimics the way users used these tags to represent songs. Based on the category, the tag distinguishing problem is then converted to a classification problem because we can rely on the class of a song to determine which tag is more accurate.

III. PROPOSED METHODOLOGY

The user is supposed to blink eyes once for requesting for text communication and twice for playing music based on the mood detected. The webcam is placed in front of the patient through which the face is detected and then the face region is extracted from the video stream using method developed by Viola and Jones. From this, face region is extracted by using Haar-Based Cascade Classifier. The Boosting Algorithms are used for best feature extraction to distinguish between face and non-face image. Extracted image consists of fixed regions of geometrical ratios for eyes, nose and mouth.

The eyes are now detected from the extracted eye region and then again the Haar-Based Cascade Classifier is used to classify both the left and right eye separately. The Histogram of oriented gradients (HOG) is used to detect an eye blink. The blinks occurred can be classified by using Support Vector Machine (SVM) Classifier which finds a decision surface two differentiate the two classes. The Yale and Cew Databases are used for purpose of testing. After identifying the appropriate blinks, we can use this for communication by forming the message. The appropriate words are also suggested for forming the message.

The face region is also used for extraction of emotions. The mood of the person is then identified, which is then processed to generate an appropriate music playlist where all the songs from the database are classified to a particular category that matches the mood to play the music.

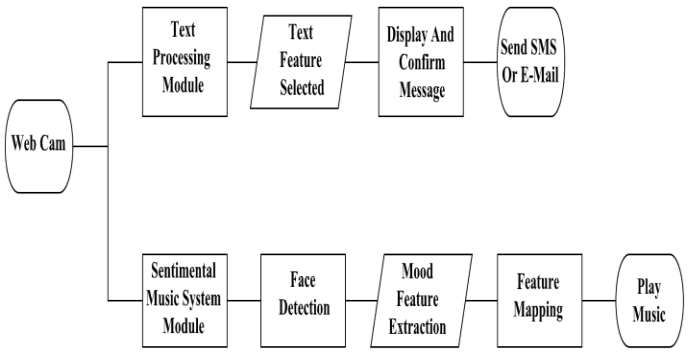

Figure 3. Interfaces split in the system

IV. CONCLUSION

The human computer interaction is important for every person who is physically impaired of this century. Blink-to-Text can be a helpful tool for those who cannot use the keyboard by their hands due to a number of reasons. It can be used by the paralyzed especially differently able persons to effectively interact with the outside environment through this application.

The implementation of other systems that are available in market require hardware like EEG headsets which is costly to use in daily life. The proposed software platform converts eye blinks to text. Every feature of the software can be controlled by eye movement. Thus, the software can be independently operated by paralyzed people. Using the software, patients can record messages, recite those messages aloud, and send the messages to others. The software can be run on any low end computer to a high end computer. The proposed work uses computer vision and Haar cascades to detect eye blinking and convert the motion into text. The program uses language modelling to predict the next words that the user might blink. The proposed software can be easily customized for each patient as well. A mood based song suggestion is also available for use by the paralysis affected patients. Thus, application provides efficient and cost-effective solution to the problem.

V. ACKNOWLEDGMENT

We are grateful to our Institution, Global Academy of Technology, with its ideals and inspirations for bringing in the quality in the project work carried out at this institute.

We earnestly thank our Principal, Dr. N. Ranapratap Reddy, and our HOD, Dr. Ganga Holi, Global Academy of Technology for facilitating a congenial academic environment in the College and for their kind support, guidance and motivation during the course of our project work. We would like to extend our sincere thanks to our parents and friends for their support.

REFERENCES

[1] Christopher and Dana Reeve Foundation: christopherreeve-stats about paralysis.
[2] MedicineNet :stats for locked in syndrome
[3] SpinalCord.com: spinal cord injury statistics
[4] IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017 on "*An Intelligent Music Player based on Emotion Recognition"*.
[5] 2012 Fourth International Conference on *"Digital Home"*.
[6] 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017 on *"An Intelligent Music Player based on Emotion Recognition".*
[7] IEEE International Conference on Computer Vision and Information Security (CVIS) *"A Novel Method for Eye Tracking and Blink Detection in video frames".*
[8] Research gate Conference Paper for *"Automatic Eye Blink Detection Using Consumer Web Cameras"*.
[9] CS 231M Project Paper Spring Quarter at Stanford University for *"Real Time Blink Detection For Burst Capture"*.
[10] 21st Computer Vision Winter Workshop, Luka Cehovin, RokMandeljc, VitomirStruc on *"Real-Time Eye Blink Detection using Facial Landmarks"*.
[11] 9th International Conference on Computational Intelligence and Communication Networks on *"Human computer interaction by eye blinking on real time"*.
[12] International Conference on Electronics, Communication and Aerospace Technology ICECA 2017 on *"face detection and tracking"*.
[13] Institute of Biomedical Engineering, National Central University on *"Implementation of an Eye-blink-based Communication Aid for People with Severe Disabilities".*
[14] Institute of Biomedical Engineering, National Central University on *"Implementation of an Eye-blink-based Communication Aid for People with Severe Disabilities"*.
[15] International Conference on *"Identifying Accuracy of Social Tags by Using Clustering Representations of Song Lyrics"*.

Authors Profile

Sameer G Mathad, Pruthvi L R, Sanjana C S and B Akash is currently pursuing Bachelor of Engineering in Information Science at Global Academy Of Technology which is affiliated to Visvesvaraya Technological University of Belagavi, India. This is one among the first papers published from all the individuals. The main research work focuses on Machine Learning Algorithms, Natural Language Processing, Server Side Scripting, Image Processing and Computer Vision. The team has just 3 years experience in this field of engineering and has undertaken many mini-projects in this duration. All the individuals would like to gain more knowledge on all these domains and do research work along with building projects which would be benefitial for the society.

Deepthi V S pursed BE in CSE from Visvesvaraya Technological University, Belagavi, India, and M.Tech in CSE from Visvesvaraya Technological University, Belagavi, India. She is currently pursuing Ph.D and currently working as Assistant Professor in Department of Information Science & Engineering, Global Academy of Technology. Her main research work focuses on Network Security, Machine Learning, Big Data Analytics, Data Mining, and IoT. She has 9 years of teaching experience and 3 years of Research Experience.

