Hand Gesture Recognition and Word Conversion for Specially Challenged and Rehabilitated People

V. Elavarasi[1], N. Parkavi[2], Dr. B. Padmanaban[3] & Dr. A. Sundar Raj[4]

[1]*Assistant Professor, Department of ECE*
[2]*PG Scholar, Department of ECE*
[3]*Professor, Department of ECE*
[4]*Associate Professor, Department of ECE*
[1,2,3,4]*E.G.S. Pillay Engineering College, Nagapattinam, India.*

Abstract

Struggles faced by the physically challenged and rehabilitated people in day to day life is unexplainable. They are having a huge difficulty in conveying their thoughts even their needs also. Though we are having improved technology, there is no efficient supporting system for differently abled people especially deaf and dumb. Hence this paper aims to provide respeech facility for the dumb people by recognizing their gesture. For this process, their gesture is recognized and converted into text and then the text get converted into speech. Here the American Sign Language System is converted into text and then speech is generated in English language from the text, all this process done by a computing device. This paper assigned unique sign representation for each and every alphabets and numeric value starting from A to Z and 0 to 9 respectively. This work uses python for programing, webcam and computer is for hardware need. OpenCV is used for capturing the gesture of the person and for training and recognition purpose machine leaning algorithm is used.

Keywords: *Hand gesture recognition, specially challenged people, American sign language, Machine learning algorithm, text to speech conversion.*

1. Introduction

In present scenario, technology is showing is tremendous growth in every sector. Our work gets simplified by technology and our difficulties are shrinking also because of its improvisation. In the similar way, we are using the advancement of technology for the betterment of differently abled people especially for deaf and dumb [1]. This paper aims to create voice for the people who cannot talk. Here, gesture of the differently abled people is taken as input for the system and it get converted into text, after the conversion of gesture to text speech get produced from the text. Gesture is nothing but a non-verbal communication medium. Information can be extracted from the gesture by recognizing the physical movement.

Figure 1. Hand recognition for capturing gesture

Gesture can be used for user interface in real-time works [2], [3] like home automation, defence, and to unlock devices. Basic steps involved in gesture recognition is, capturing of the sign, processing by software or algorithms and comparing it will already predetermined gesture in the library. Once the input gesture get matched with the gesture in library, then the system will follow the command connected with the gesture. In this method the process starts from capturing the gesture by using web camera. For the first step in processing, histogram of the input hand is obtained. Then the gesture is captured for training and there is a separate training dataset is available. Then the recognition of the gesture takes place for converting the gesture to text. This is done by matching the input gesture with gestures already stored in the database. And finally, the text get converted into speech. For text to speech conversion, simple text to speech synthesizer is used.

2. Proposed method

Hand gesture is the most efficient way of interaction for differently abled people. By capturing the hand gesture [4] and developing a technology which will recognize the gesture and give them a speech format will definitely help the differently abled people in their day to day life. The overall working of this method can be represented in developer and user point of view. First I case of developer, this process starts with capturing the image through the web camera. Then the original image of the hand is converted into histogram format. Various algorithms like HSV-Skin Detection, Large Blob Detection, and Contour Detection are used for recognition of hand gesture. And for speech production Speech and Text Synthesizer is used. At first the video is captured and converted into frames in RGB format then, histogram of the hand image is obtained. After this process, skin colour of the hand is identified by using HSV-Skin detection algorithm. And the processing of hand image involves few more steps. Overall steps involved in this hand gesture to speech conversion is shown in figure 2.

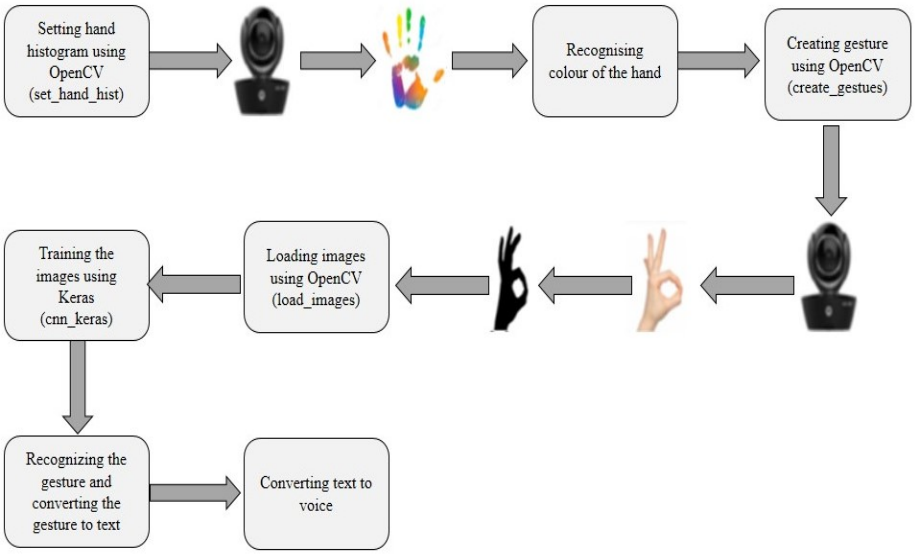

Figure 2. Overall architecture of proposed method

After skin colour detection, grayscale image is obtained from the previous step by using Blob detection algorithm for accurate detection of skin. Then the frames or image is taken for training. For training, our proposed method will take 1200 images for a single

gestures. Then the images stored in training dataset and used for training module. Keras and Convolution Neural Network (CNN) [5] is used for training process. Only after completing the training of dataset, the system become capable of recognizing hand gestures. Recognition of hand gesture takes place in developer and user side which is represented in figure 3. Similar to CNN, various algorithms like Least Mean Square algorithm are used for eliminating dominant harmonics from line current [6], input current harmonics suppression [7].

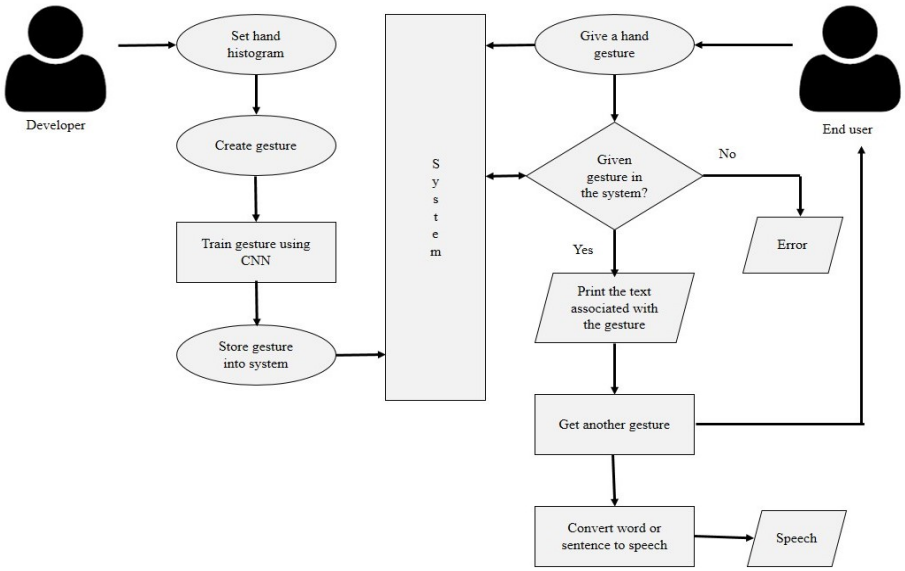

Figure 3. Process involved in developer and user side of proposed system

In user side, first the hand gesture is captured through the webcam and then as same in the developer side, it get converted into histogram image. Then the system will search for the match in the dataset. Once it got the correct match of the input hand gesture it will give the corresponding text for that gesture as output. Finally the text will get converted into speech by using Google text to speech API [8].

3. Implementation of proposed method

In real-time implementation of the proposed method shown in figure 4, the process starts with capturing the hand image of the differently abled people or user [9]. Hand image is nothing but the gesture shown by the differently abled people.

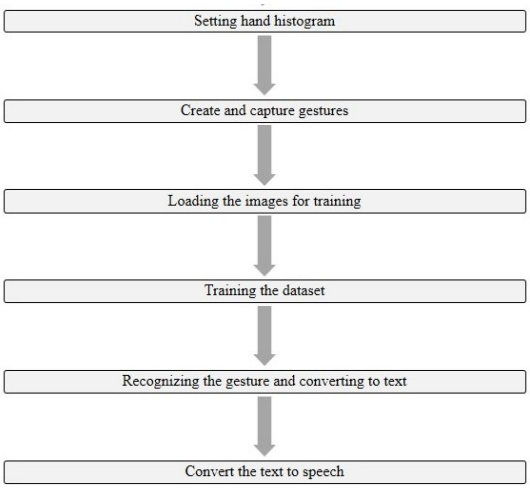

Figure 4. Steps in implementation of proposed method

First, the raw image is converted into histogram. Histogram is nothing by a graphical representation of tonal distribution in digital image. Here each tonal values get represented by number of pixels. In histogram representation, pixel are given either high or low value. Histogram representation involves plotting the number of pixels in the image (vertical axis) with a particular brightness value (horizontal axis). This process mainly helps in edge detection and image segmentation. In real-time process, the user has to place their hands in front of the camera by covering the 50 squares which present in the input screen shown in figure 5. Inorder to detect their skin colour the user has to press the 'c' option. Then the monitor will display the pixels similar to the skin colour of the user in the 'thres' window.

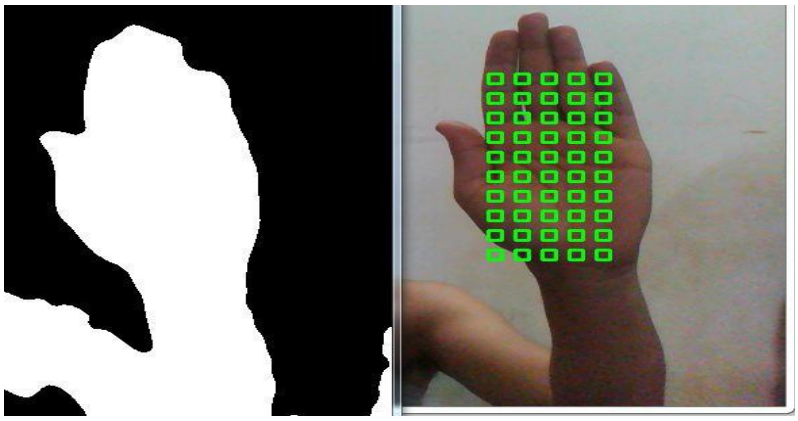

Figure 5. Setting hand histogram of input image

Next step in implementation of this process is, gesture capturing and creation. Here the gesture from the input is captured for the training dataset. In this process, the user has to enter the gesture number and name. OpenCV window will open for capturing the gesture. The system will continuously capture the gesture and the user can move their hands a little bit for capturing different looks of the gesture. System will kept count of the number of gesture image captured. This process can be paused and resumed by pressing 'c'. Once the count reaches 1200, window will close automatically.

The captures 1200 images of each gesture is then moved to gestures folder for training purpose. This process of moving the image for training is known as loading and it will create four files named test_images, test_labels, train_images, train_labels [10]. The

stored dataset are then taken to training. For training, this system uses keras and CNN based method. In CNN, single input image is processed in multiple stages for better output. By using keras, the neural network model for processing is converted into sequential model. Sequential training process involves large number of iteration for getting various discriminative features and parameters among the dataset. Convolution 2-D can be used for image and convolution 3-D can be used for processing video [11].

After completion of training, the model become able to recognize the gesture. Recognition process is done by considering the position of fingers and palm. And the captured input gesture is first resized to 224x224 pixel and then compared with the already stored gestures in the dataset. If the module found the matching gesture for the input, then the corresponding text stored in the dataset will get displayed. Because this method has unique gesture representation for each and every alphabets and numbers. Then the obtained text is converted into speech. For text to speech conversion, our proposed method utilizes Google Text to Speech API(GTTS). At last, the word for the given sign language of the user is pronounced as a word.

4. Result

Since the method proposed in this paper is having real-time application, it is very necessary to maintain the accuracy along with very less delay. After capturing the image of gesture, the process of skin colour or object colour detection shown in figure 6 and obtaining histogram representation shown in figure 7 is done. Once the gesture is captured in image, it will get compared with the 1200 images of each gesture stored in database.

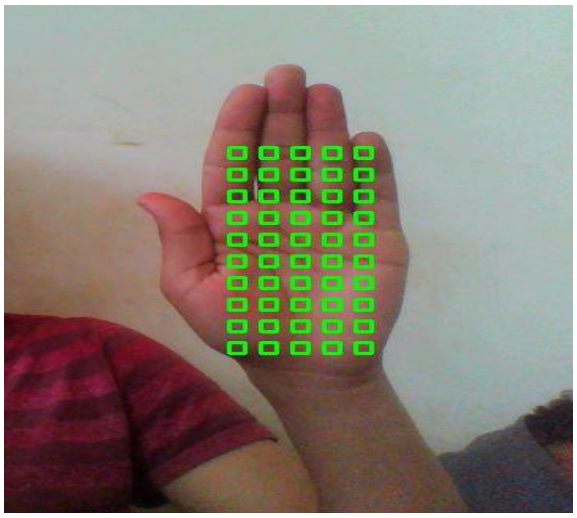

Figure 6. Colour detection

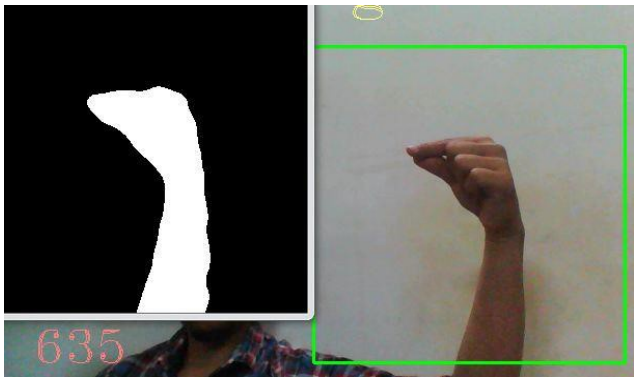

Figure 7. Creating gesture and its histogram

The captured images of the gesture is initially used for training. For training our module continuously capture 1200 images of a single gesture and it is loaded for training shown in figure 8.

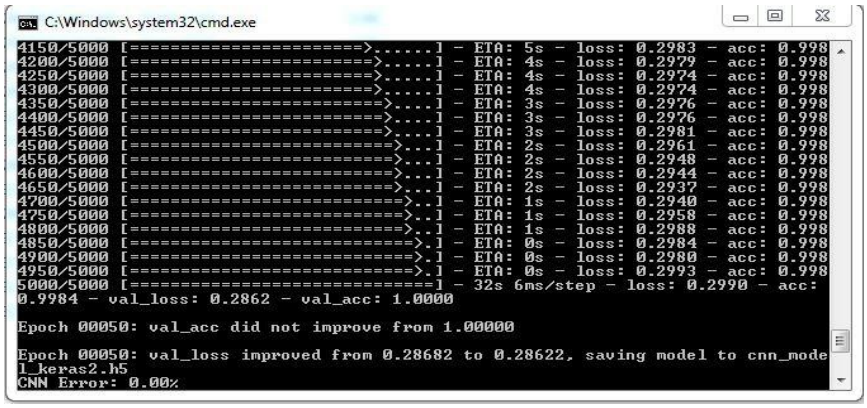

Figure 8. Training the gesture for recognition

Once the training is successfully done, our module is ready to get implemented for real-time application. During the working of our module, if the user (differently abled person) shows the sign language (gesture) it get captured and processed as shown in figure 6 and 7. Then the captured gesture get compared with the dataset for correct match. If the system found the correct match, it will display the corresponding text shown in figure 9. And the converted text is then converted into speech using GTTS.

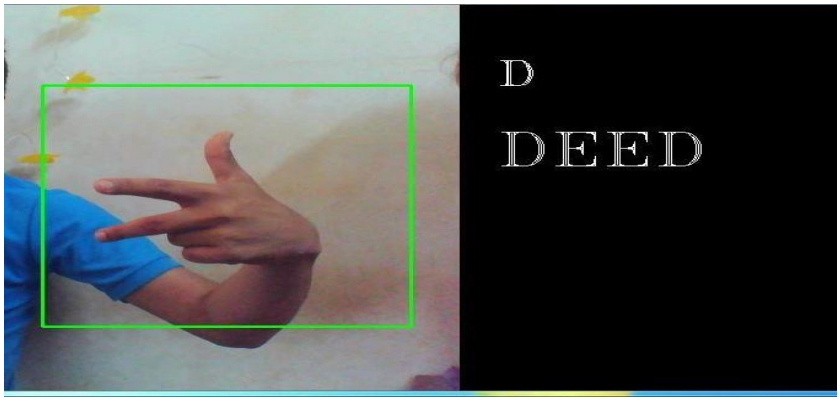

Fig. 9 Gesture to text

5. Conclusion

Differently abled people are facing huge struggle to communicate with other person. Hence this paper aims to add some betterment in their life especially for deaf, dumb and also rehabilitated people. Since this module gives artificial voice to the vocally impaired people, they become more independent. Here we are using the hand gestures of the user and convert them to speech through various process. The conversion process of hand gesture to speech is done in high speed with good accuracy. In future, this module can be implemented in smart phones so that it will become accessible to every vocally impaired people.

REFERENCES

[1] Vishmita Yashwant Shetty, Vinayak Bhupendra Rai, "Sixth Sense Technology", Journal of Nanosci. Nanotechnology, vol. 3, no. 12, **(2014)**, pp. 1067-1073.

[2] Jian Wu, Lu Sun, Roozbeh Jafari, "A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors", IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 5, **(2016)**, pp. 1280-1290.

[3] Hidetoshi Ishiyama, Shuichi Kurabayashi, "Monochrome glove: A robust real-time hand gesture recognition method by using a fabric glove with design of structured markers", Proceedings of the 2016 IEEE Virtual Reality (VR), SC, USA, **(2016)** March 19-23.

[4] Yu Wu, Dai Jiang, Jifang Duan, Xiao Liu, Richard Bayford, Andreas Demosthenous, "Towards a High Accuracy Wearable Hand Gesture Recognition System Using EIT", Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, **(2018)** May 27-30.

[5] Marouane Benmoussa, Abdelhak Mahmoudi, "Machine learning for hand gesture recognition using bag-of-words", Proceedings of the 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, **(2018)** April 2-4.

[6] Mohan.V, Jeevananthan.S, Raja.J, "An on-line Adaptive Filtering for Selective Elimination of Dominant Harmonics from Line Currents of a VSI fed Drive using Recursive Least Square Algorithm", Proceedings of the IEEE International Conference on Advances in Engineering, Science and Management (ICAESM-12), Nagapattinam, Tamil Nadu, India, **(2012)** March 30-31.

[7] Mohan.V, Raja.J, Venkatesh.C.H, Jeevananthan.S, "Adaptive LMS algorithm based Input Current Harmonic Suppression in AC-DC-AC Drive with either Constant or Pulsed DC Link", Proceedings of 46th International Universities Power Engineering Conference (UPEC 2011), Soest, Germany, **(2011)** September 5-8.

[8] Neela Harish ; S. Poonguzhali, "Design and development of hand gesture recognition system for speech impaired people", Proceedings of 2015 International Conference on Industrial Instrumentation and Control (ICIC), Pune, India, **(2015)** May 28-30.

[9] Bhavya Shahi, Mansi Goyal, K.V. Prema, N.V. Subba Reddy, "A novel approach to human gesture recognition", Proceedings of 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bangalore, India, **(2017)** August 17-19.

[10] R. H. Goudar ; Shwetha S Kulloli, "A effective communication solution for the hearing impaired persons: A novel approach using gesture and sentence formation", Proceedings of 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bangalore, India, **(2017)** August 17-19.

[11] Tushar Chouhan, Ankit Panse, Anvesh Kumar Voona, S. M. Sameer, "Smart glove with gesture recognition ability for the hearing and speech impaired", Proceedings of 2014 IEEE Global Humanitarian Technology Conference - South Asia Satellite (GHTC-SAS), Trivandrum, India, **(2014)** September 26-27.