Proceedings of the International Conference on Intelligent Computing and Communication for Smart World (I2C2SW-2018)
14-15 December 2018

Portable Communication Aid for Specially Challenged : Conversion of Hand Gestures into Voice and ViceVersa

T. MeeraDevi
Department of ECE
Kongu Engineering College, Perundurai
Erode, TamilNadu
meeradevi@kongu.ac.in

K.M. Sharavana Raju
Department of CSE
College of Computer Science and information Technology
Jazan University, Jazan, Kingdom of Saudi Arabia
smaruthachalam@jazanu.edu.sa

***Abstract*— Communications between a normal person and with the person having hearing loss and dumb have constantly been a tough assignment. The work is to develop a portable device for the disabled people those who are not able to communicate with the normal persons properly. The technology development presents a solution to build up a sign language conversion system to support the individual with hearing loss and mute people. The core idea is to build up a real time embedded product for the disabled persons without handheld gloves to assist their announcement in efficient way. Similarly the speech communication by normal persons will be converted into gestures for the disabled persons for their better understanding. This device will act as a two way communication device between normal and disabled persons.**

Keywords—Communication aid, Sign language, Hand Gesture

I. INTRODUCTION

More in recent times urbanized embedded applications are altering our standard of living in an elegant way. Sign language is a communicative and normal way for communication among common and the human being with hearing loss and mute [3]. The objective of the sign language translation system is to decode the usual sign language into speech and to create trouble-free contact with the dumb persons. This system is developed in order to improve the way of life of the dumb person..The physical gesticulation statement comprise of hand gestures that suggest particular meaning, the non physical is head movement, facial appearance, body orientation and position.

The most important uniqueness of a person hand gesticulation can be reviewed by their energetic features. To exploit hand gesticulation as a mode of interface, it is obligatory to investigate the actions for each of the gesticulation and finally to pull out the whole understanding of the gesticulation. broadly ranging from restoring the customary mouse as a point tool to virtual reality and communication with the hearing loss patients.

The entire processcontains the subsequent concerns.

- recognizingthe features distinguishing hand gesticulation.
- Train the collected gesticulation using Neural Network
- Separating the hand movement pattern from a continuous recording of gestures
- Low-level understanding for the feature patterns comprises the gestural segment.
- Understand a whole interpretation incorporating the concurrent features.

II. HAND GESTURE TO VOICE

Hand gesture conversion to voice involves images of live hand gestures taken lively with which training datasets are developed and with this dataset the gestures are recognized. The entire process is shown in the figure 1

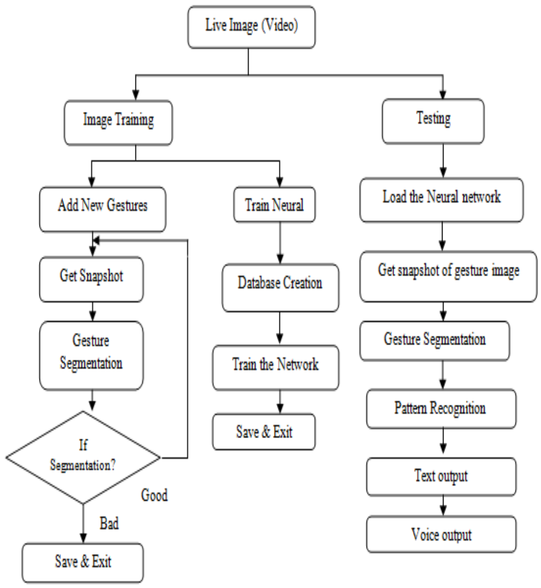

Fig. 1. Hand Gesture to Voice

The translatorconsists of training and testing phases.

A. Training phase

- The process involved in the training phase is given in the following steps.
- Training dataset is created by capturing different hand gestures through high resolution camera. For each gesture minimum of ten images are captured.
- The captured images are de-noised using median filter.
- For better training, hand gesture alone is segmented by removing the background information. Skin color based segmentation ie. RGB and YCbCr color

transformation is used for segmenting the hand gesture alone.

- Feed forward neural networks with 200 hidden layers is used to train the dataset.
- In this prototype design a minimum of 10 gestures were recognized, so the network consists of 10 outputs.
- Number of Training epochs is 58 epochs with 0.84sec.

B. *Testing phase*

In this phase the trainedsystem in the previous phase is used reconditioning the gestures. In real time gesture, a dumb person live gesture is captured. This captured image has gone through the filtering and segmentation process and then it is sent to the classifier network.

If the gesture matches with dataset by the testing phase then text output of the corresponding gesture is displayed on the screen of the handheld device. Datasets considered in this project is shown in table 1.

TABLE I. DATASET TABLE

Gesture type	No of samples
Numerals 0 -9	20 each
Alphabets a-z	17 each
Alphabets A-Z	20 each
Okay	16
Stop	12
Rest Room	15

As in the given table 1, 20 different images for training the number 'one' is shown in figure 2.

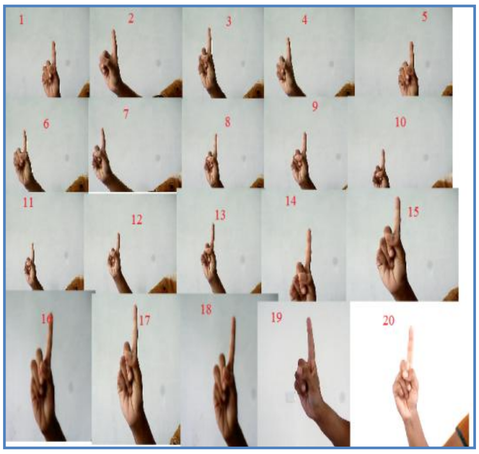

Fig. 2. Training Dataset for 'ONE'

The result of Skin color based segmentation as shown in figure 3.

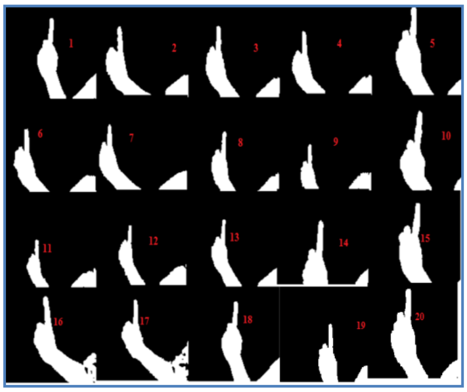

Fig. 3. Segmented Dataset for 'ONE'

C. *Research Results*

The nntraintool shown in figure 4specifies the number of input, number of output and number of hidden layer size, number of andits performance measures.

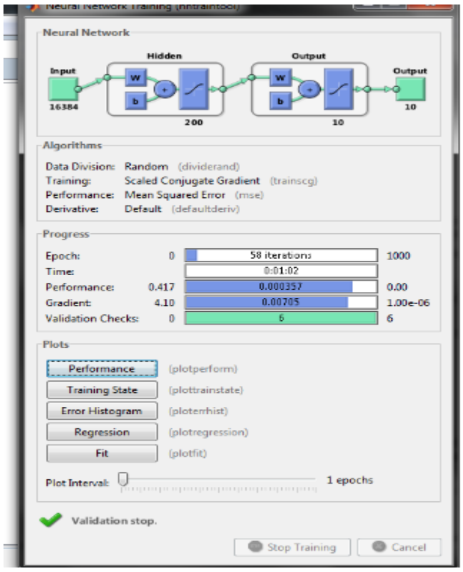

Fig. 4. Training Parameters

For training the dataset shown in table 1, the proposed system undergone 58 iterations at 0.84 secs. The confusion Matrix for the firstphase is shown in table 2. The overall percentage of the firstphase is about 95% .

TABLE II. CONFUSION MATRIX FOR THE FIRST PHASE

Data set	True positive Value (%)	True negative Value(%)
Numerals (0 -9)	94.4	5.6
Alphabets (a-z)	95	5.0
Alphabets (A-Z)	95.5	4.5
Okay	94.7	5.3
Stop	100	0
Rest Room	100	0

The trained system performs pattern recognition after the segmentation is done. Here the text output 'OKAY is displayed in the screen for the hand gesture OKAY as in figure 5

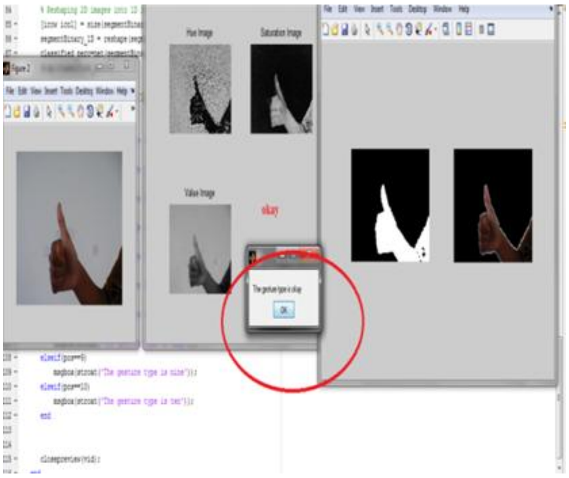

Fig. 5. Text Output for 'okay'

Similarly the text outputsZERO' and 'STOPare displayed on the screen '' ashown in figures 6 and 7.

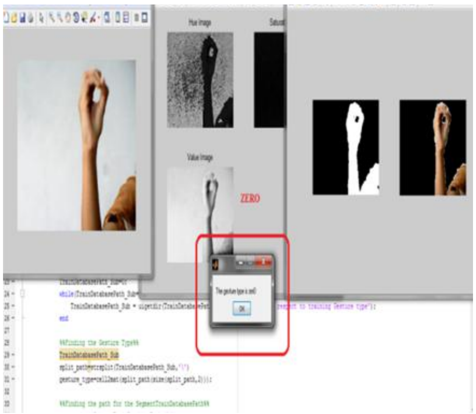

Fig. 6. Text Output for 'zero'

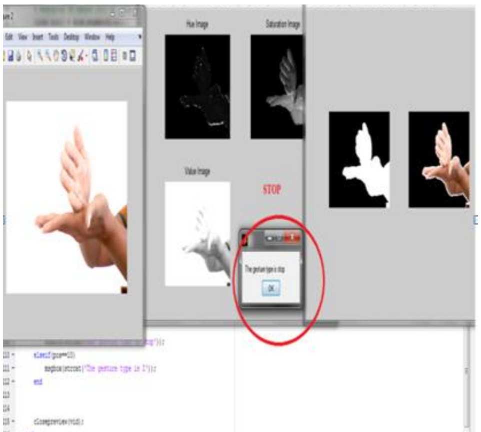

Fig. 7. Text Output for 'stop'

From the text output from the testing phase is converted into voice using text to voice converter with a hardware setup and is shown in the figures 8 and 8 for the outputs 'FIVE' and 'REST ROOM'.

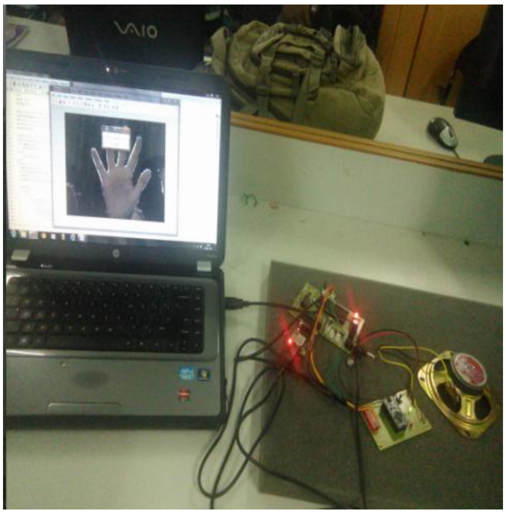

Fig. 8. Voice output for 'FIVE

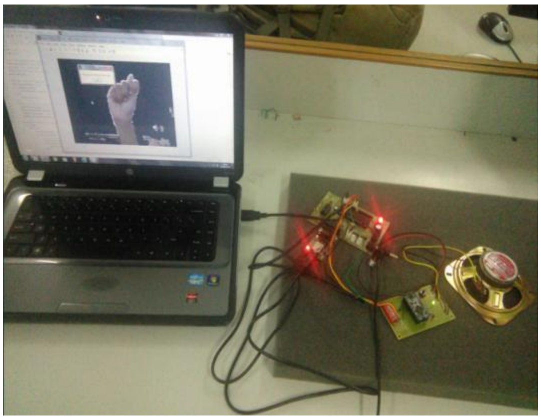

Fig. 9. Voice output for 'REST ROOM'

III. SPEECH TO GESTURE CONVERSION

In the opposite process, the normal person voice has to be converted into gesture for the understanding of disabled persons. Speech recognition involves voice recognition and converting it into text and gesture images. This entire process in given in the flowchart shown in figure 10.

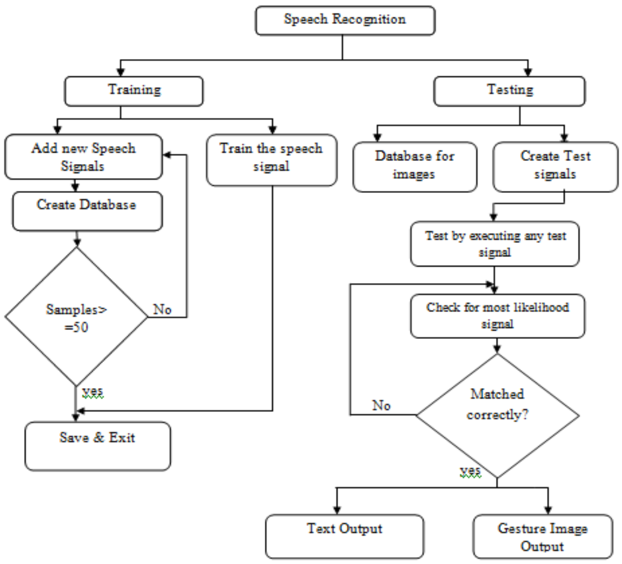

Fig. 10. Voice to Hand Gesture

Speech signal of a normal person is collected with minimum of 50 speech samples in each case. These speech signals are trained for detection using LPC algorithm. A test case is created for testing the speech signals. Mahalobian distance is used along with the LPC algorithm to check for most likelihood speech signal. After the testing is being done the corresponding text and gesture images are displayed. Speech recognition for ‘DAD’ is shown in the figure 11.

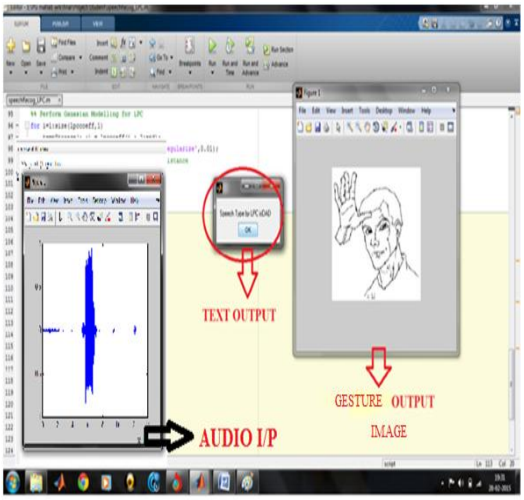

Fig. 11. Voice input ‘DAD’ to Hand Gesture

The input speech is recognized and the corresponding text and the gesture images are displayed as shown in figure 12 which is speech recognition for ‘THANK YOU’

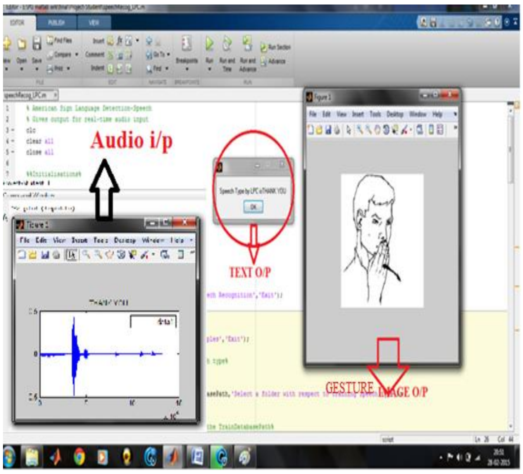

Fig. 12. Voice ‘THANK YOU’ toHand gesture

The LPC algorithm is a prediction based algorithm used for speech recognition. The speech recognition for ‘PLEASE’ is shown in figure 13

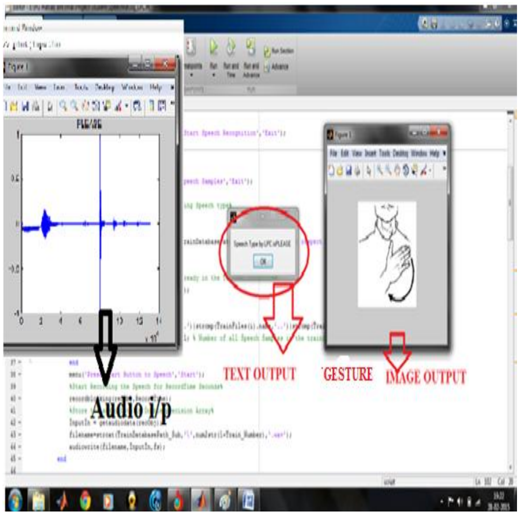

Fig. 13. Speech recognition of ‘PLEASE’

Speech recognition for ‘REST ROOM’ using LPC algorithm (which is a prediction based one) is shown in figure 14. The input speech signal from the test set is executed and similarity is found and corresponding text and gesture image is displayed

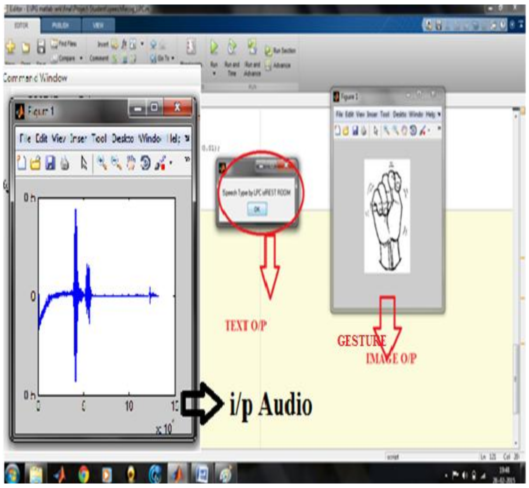

Fig. 14. Speech recognition of ‘REST ROOM’

The speech recognition by LPC algorithm for ‘COFFEE’ is shown in figure 15 in which a text and a gesture image is produced for the input speech signal.

The input speech signal from the test set is executed and similarity is found and corresponding text and gesture image is displayed. The speech recognition for ‘HELP’ is shown in the figure 16

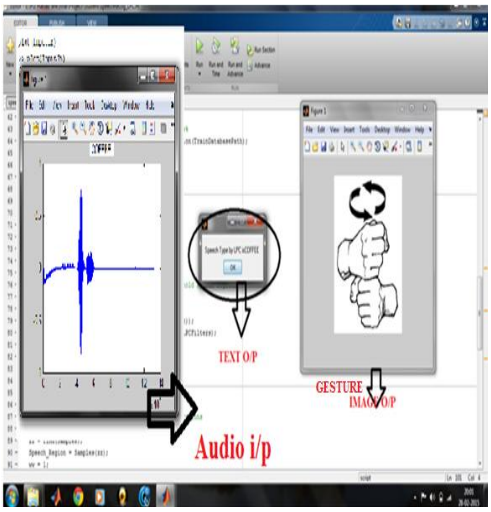

Fig. 15. Speech recognition of 'COFFEE'

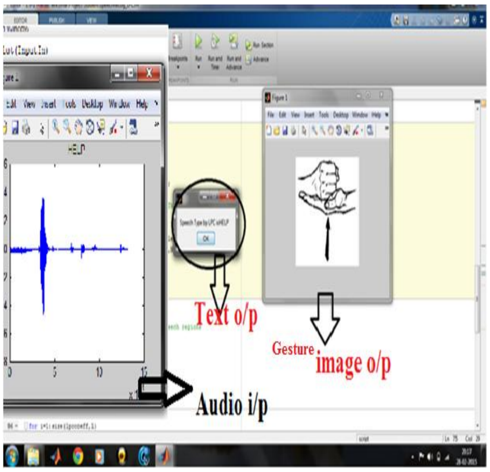

Fig. 16. Speech recognition of 'HELP'

The input speech is recognized and the corresponding text and the gesture images are displayed as shown in figure 17 which is speech recognition for 'SORRY' by using a prediction based algorithm.

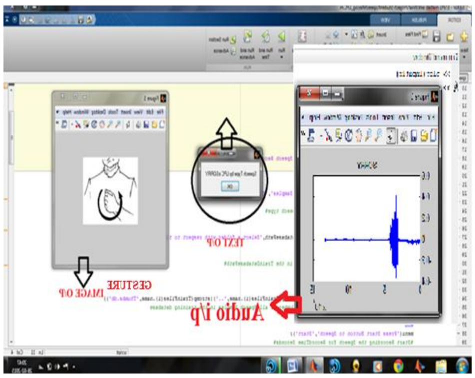

Fig. 17. Speech recognition of 'SORRY'

IV. Conclusionand Future work

This system has been developed for set of different words and its voice and for different gestures. This proposed will be useful for the normal persons to communicate with differently abled persons (The person with hearing loss with mute) and vice versa. In future this system may be extended for variety of words and voice inputs and has to be developed as a handheld device as a guide for the specially challenged people and enable them to communicate with all without any gap.

References

[1] Aditi Kalsh, N.S. Garewal , "Sign Language Recognition for Deaf & Dumb", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 9, September 2013.

[2] Beifang Yi, Frederick C. Harris Jr., Ling Wang, and Yusong Yan, "Real-Time Natural Hand Gestures," IEEE journal on Computing in Science and Engineering, 2005

[3] A. Chaudhary, J.L. Raheja, K. Das, S. Raheja, "A Survey on Hand Gesture Recognition in context of Soft Computing", Published as Book Chapter in "Advanced Computing" CCIS, Springer Berlin Heidelberg, Vol. 133, 2011, pp. 46-55

[4] Ibrahim, P., and Srinivasa, R. "Automated speech recognition approach to continuous cue symbols generation", International journal of power control signal and computation, Vol. 18, No. 8,pp. 434-520,2007.

[5] J.L. Raheja, A. Chaudhary, S. Maheshwari, "Automatic Gesture Pointing Location Detection", Optik: International Journal for Light and Electron Optics, Elsevier, Vol. 125, Issue 3, 2014, pp. 993-996.

[6] Miada A. Almasre, Hana Al-Nuaim, "A Real-Time Letter Recognition Model for Arabic Sign Language Using Kinect and Leap Motion Controller v2" International Journal of Advanced Engineering, Management and Science (IJAEMS), Vol-2, Issue-5,ISSN: 24541311,May 2016

[7] Mostafa Karabasi, Zeeshan Bhatti, Asadullah Shah, "A model for Real-time Recognition and Textual Representation of Malaysian Sign Language through Image Processing," International Conference on Advanced Computer Science Applications and Technologies, 2013

[8] Shiguo Lian, Wei Hu, Kai Wang, "Automatic User State Recognition for Hand Gesture Based Low-Cost Television Control System," IEEE Transactions on Consumer Electronics, Vol. 60, No. 1, February 2014

[9] Soo-Chang Pei, and Yu-Ying Wang, "Census-based Vision for Auditory Depth Images and Speech Navigation of Visually Impaired Users," IEEE Transactions on Consumer Electronics, Vol. 57, No. 4, November 2011

[10] Yuan Yao, and Yun Fu, "Contour Model based Hand-Gesture Recognition Using Kinect Sensor," IEEE Transactions on Circuits and Systems for Video Technology, 2013

[11] Zhengmao, Z., Prashan, R., Monaragala, N., Malin, P. "Dynamic hand gesture recognition system using moment invariants", IEEE Transactions on neural networking and computing, Vol. 21, No. 1, pp. 1034-1320, 2010.