

Gesture and Speech Recognition through Programming Interface

Chigurala Shruthi¹ and B. Sreenivasu²

M Tech Scholar¹ and Associate Professor², Department of Electronics and Communication Engineering
Sreyas Institute Of Engineering & Technology, Beside Indu Aranya, GSI, Bandlaguda, Nagole, Hyderabad,
Telangan

ABSTRACT

Sign language is the preferred methodology of communication among the deaf and the hearing impaired people all over the world. Now-a-days text and sign language processing are the most important terms in digital image and text processing industry. The sign language recognition is the most popular research area involving computer vision, pattern recognition and image processing. This system is intended to design the effective system, which can provide the ultimate support to deaf-and-dumb candidates for their day-to-day life activities. The sign language processing helps us to capture the text from web camera and pronounce that text using voice based applications. In this approach the constant text format the user need to follow such as the font type of Arial or Calibri with the size of 48 to 50 points. Because this sign language processing is a Natural Language Processing scheme, which can acquire text from real-world scenarios, so that the text must be in standard format. The voice gathering and text conversion system, allows the user to work with mobile application, which converts the respective user's voice into text using goggle API. For all the entire system provides good and effective support to deaf-and-dumb persons without any interruptions

Keywords: Sign Language Recognition, Application of Sign Language, Data Input and Voice Output, Deaf and Dum, API.

I. INTRODUCTION

There are number of sign languages across the world and it is unique for every country based on their spoken languages. Considering the languages, in which there are Srilankan, Bangladesh, Japan, Russian, American, British, Indian sign languages. In all these sign language as well International Sign language is English, American sign language is used by almost all the people across. There are a lot of variance between American and British signing as shown in fig 1. And almost all uses ASL is used across which is standards for all languages like Indian sign Language.



Fig 1: Contrast between American signing and British signing

DISABILITY TYPES

Generally there are various disabilities which differ from person to person. As the scenario considered in DEAF, DUMB, BLIND people can't express their communication to rest of the world in which they can't hear, speak and see respectively. Two communities people are helpless in both speaking skills and listening skills in fact no one can help them. One and only simple thing that separates them and the normal people is promulgation. Supposing that may be a path for disable persons and the normal people to promulgate, the deaf-mute human beings can without difficulty stay as a ordinary individual. More over the simplest contrivance for them to speak only thru Gesture language .While Gesture may be very vital to deaf-mutes humans, for speaking both with normal human beings and with themselves, continues to be getting tiny interest from the normal people. Because ordinary humans, have a tendency to reject the significance of

gesticulation, merely cherished persons who are deaf-mute. Unique solution to talk with the deaf-mute humans is by way of the usage of the offerings of Gesture interpreter. But the usage of gesture interpreter may be highly-priced. So intelligent answer is needed so that the deaf-mute and everyday humans can communicate conventionally. consequently, analyst need to discover a manner for the deaf-mute human beings a good way to communicate without difficulty with regular person. The leap forward for that is the Gesture Recognition System. Along with deaf -mute persons, people who are blind by birth or by accidental cannot continue their fulfillments like reading books, newspapers or education etc. So considering their conditions with the help of Optical Character Reader they can read and pertain knowledge.

MOTIVATION

The main motivation of the project are by seeing the disable persons all around the society and families. The deaf mute community people are not able to understand the words spoken by normal persons and cant talk with the normal persons. Some of the people are as shown in below fig 2 ,the images are taken from a statistical survey of disabled persons in India in year 2016. To improve communication among all persons i have chosen this project.



Fig 2: People suffering with disability (statistical survey 2016)

II.OBJECTIVES

The main objectives of the system is:

- To provide efficient two way communication by using English language.
- To develop an interactive atmosphere between normal and disable people.
- Communication can be done with the Android mobile which everyone is using.

III.PROPOSED SYSTEM

From the analysis of literature survey and problem statement, we have proposed a new methodology for Sign language recognition with the hand gestures. The proposed system consists of software and hardware. This consists of following main steps ,they are

1. Capturing
2. Recognition
3. Segmentation
4. Extraction of Features
5. Classifier

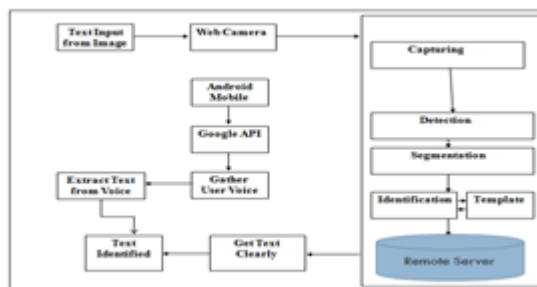


Fig3: Proposed System Block Diagram

CAPTURING

The images shown in front of the web camera of laptop are captured and fed as input to the system. The images may include words, hand gestures, alphabets considering the all background data as shown in fig 4.a,b,c images Then the image is further enhanced for next stage



Fig 4: Number of sample images showing in front of camera for capturing

RECOGNITION

The ability of computer to recognize that the photograph contains hand gesture and tells where it is located. For the detection we use ANN(Artificial Neural Networks).An artificial neural system is an interconnected gathering of nodes.Here every round about hub speaks to a counterfeit neuron and a arrow speaks to an association from the yield of one artificial neuron to the contribution of another r as shown in fig 5

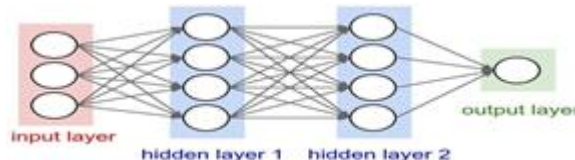


Fig 5:Artificial Neural Networks

In hand gesture recognition, by using ANN, only the required hand gesture is taken as shown in fig 6 (b) by eliminating all the background data, as shown in fig 6 (a) by passing through the hidden layers of ANN .



Fig 6:Detection image from original image

SEGMENTATION

Segmentation is useful and crucial in image recognition system because it can extract regions of interest from images which are meaningful for further processing. In the field of example recognition,machine learning,image preparing and PC vision picture

segmentation is a fundamental task.For the segmentation we are using automatic image thresholding process called **Otsu's method** which gives us a single threshold image by seperating background data and required data.

OTSU'S METHOD

This algorithm mainlysearches for the threshold that minimises the variance,which shows by taking two variables weighted variances

$$\omega_{\sigma}^2(p) = \sigma_0(p)\omega_0^2(p) + \sigma_1(p)\omega_1^2(p)$$

where σ_0 and σ_1 are the probabilities seperated by threshold p and ω_0^2 and ω_1^2 are the variances of the class.

The class probability is computed as M bins of histogram

$$\sigma_0(p) = \sum_{k=0}^{p-1} L(i)$$

$$\sigma_1(p) = \sum_{k=p}^{M-1} L(i)$$

which expresses in way that class probability L and class mean means α where the class means $\alpha_0(p), \alpha_1(p)$ and α_p are

$$\sigma_0(p) = \frac{\sum_{k=0}^{p-1} k p(k)}{\sigma_0(p)}$$

$$\sigma_1(p) = \frac{\sum_{k=p}^{M-1} k p(k)}{\sigma_1(p)}$$

$$\sigma_p = \sum_{k=0}^{p-1} k p(k)$$

To verify these equation,we follow

$$\sigma_0 \alpha_0 + \sigma_1 \alpha_1 = \sigma_p$$

$$\sigma_0 + \sigma_1 = \sigma_p$$

By computing variance and mean,yields effectiveness of algorithm.

The sample image after applying Otsu method is as shown below



Fig 7:Threshold image after segmentation

EXTRACTION OF FEATURES

It is a feature that reduces dimensionality and represents only effective parts of the image from a lumpy full of image. Feature extraction can be detected by points on the hand which gives the data required. From the threshold image of segmentation part the contour or instance points on hand tells us sign gesture. For this we are using Haar like Features.

HAAR WAVELET TRANSFORM

To face the competitive world Michael Jones has proposed a perfect detection of object using Haar transform. There are many transforms, but we use Rectangular Haar transform.

A simple Rectangular Haar feature defines contrast of aggregate of pixels of regions inside the square shape which is at any position and scale inside the actual picture..This conversion is called 2rectangularframe. As it is hand gesture recognition 2-rectangular feature sufficient transformation and to extract features with the help of edge features,line features and center surrond features as shown in below fig 8

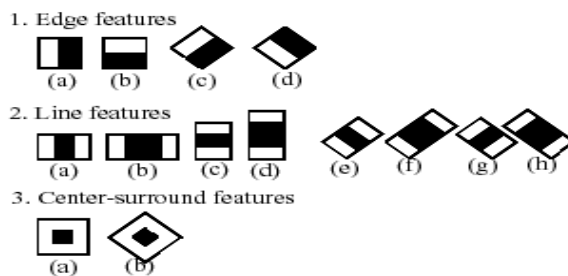


Fig 8 : 2-Rectangular features

To describe the feature extraction we use Haar wavelet transform which is simple to understand. This is not continuous and not discrete, but it can detect any transitions when applied to particular image.

The Haars wavelet is as shown

$$\phi(x) = \begin{cases} 1 & 0 < x < 1/2 \\ 0 & \text{otherwise} \\ -1 & 1/2 < x < 1 \end{cases}$$

Its scaling function is described as

$$\rho(x) = \begin{cases} 1 & 0 < x < 1, \\ 0 & \text{otherwise} \end{cases}$$

The haar transform uses 2*2 matrix i.e.,

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Using wavelet transform one can transform any type of sequence like

$(a_0, a_1, \dots, a_{2p}, a_{2p+1})$ into two component vector sequence

$((a_0, a_1), \dots, (a_{2p}, a_{2p+1}))$.

H_2 multiplies with each vector component and we obtain result $((f_0p_0), (f_1p_1), \dots, (f_m p_m))$

where f is sequences of averages and p is the details of image.

From fig 9 it can be understood that by applying Haar Wavelet Transform ,the red marks are shown in image which tells the required portion is captured in rectangular form of hand image



Fig 9 : Feature extraction of image

CLASSIFIER

Highlights extricated during the element extraction stage are given as a contribution to order step. An assortment of arrangement systems are accessible which are valuable for perceiving the motions. Arrangement step comprises of two stages: preparing stage and testing phase. For the Classification we use Support Vector Machine for recognizing the words that matches the database.

SUPPORT VECTOR MACHINE

SVM is a managed learning procedure. It was presented by Vapnik. In this strategy, each data thing is addressed as a point in a n-dimensional space with the estimation of every component being the estimation of a particular co-ordinate. The arrangement is then performed by finding the hyperplane that separates the two classes great. SVM builds up a hyper plane which has the greatest partition to the nearest planning data reasons for any class, since when in doubt the bigger the edge, the lower the speculation mistake of the classifier .The target of SVM is to make a model which predicts the class name of data cases in the testing set successfully. It uses reinforce vectors (SVs) part to outline data from information space to a high-dimensional component space which urges the issue to be dealt with fit as a fiddle . SVM is computationally costly and in this way it runs moderate. SVM produce precise outcomes. It is vigorous to noise.

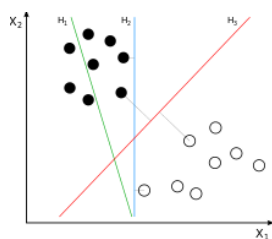


Fig 10:SVM training with the samples

In fig 10, H1 doesnot separate classes,H2 separates somewhat but H3 completely will give partition. The data available near margins are called support vectors. The sample image used in my project is as shown fig 11



Fig 11: Data received from hand gesture app

IMPLEMENTATION OF HARDWARE

The hard ware used is Raspberrian pi model-B with 1GB camera which is very demandable in present days as per cost and technology.Its features are very feasible to all age groups and easy to operate with keyboard which is used in python language platform.

The Raspberry pi model is as shown in fig 12



Fig 12 :Raspberry pi model-B

Edit with idle 3.1(64 bit) the command prompt will be opened as shown in fig 13

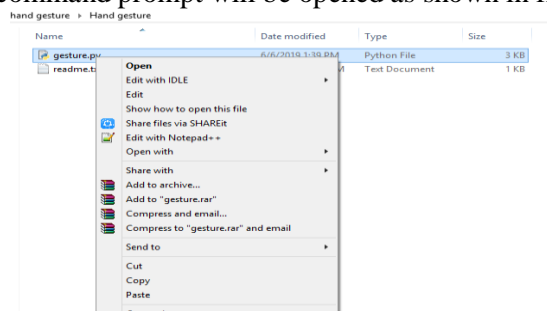


Fig 13:Opening of gesture file

After opening with idle the program is opened as shown in fig 14

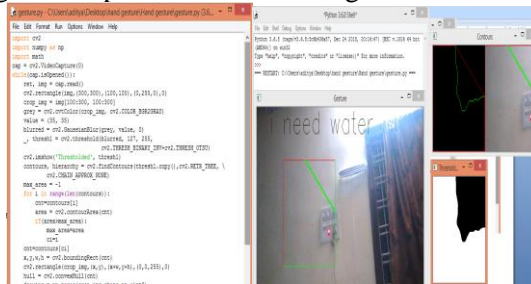


Fig14:Program for hand gestures.

IV.RESULTS AND DISCUSSIONS

The proposed system is implemented both software and hardware.

4.1 SOFTWARE

The Optical Character Recognition is based application recognized with the c# code at the backend process which is executed in Microsoft Visual Studio. The characters of the words are recognized accurately and spelt accordingly. To test the words we have nearly taken 50-100 words and tested to see the recognition and pronunciation of words. Almost all the words are recognized accurately.

REGISTRATION

For the characters to be recognized in OCR ,it needs registration and authentication of user as shown in below figure 15

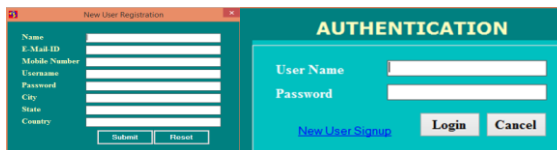


Fig 15: Authentication of user

SCANNED DATA

The fig 16shows the input data which is to be scanned which are written in Bold letters with large font.The input data can be taken anything as alphabets,words,and sentences upto 10 words.

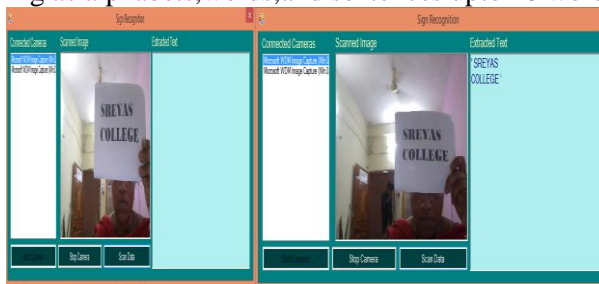


Fig 16:Input 1and scanned output

The output fig4.4 is detected correctly on right side which is exactly same as input image



Fig 17 : Input 2 and scanned output

The output in fig 17 shows correct data along with two un necessary letters in right hand side which is not included in input left side. But it pronounced correctly as input.



Fig18 : Input 3 and scanned data

The data of input fig18 is scanned and detected correctly with no issues which will be helpful for blind also to read words instead of signaling.

4.2 HARDWARE

The hand gesture recognition is in Raspberrian operating system with python based programming in open cv application. The hand gesture for five fingers are detected to fulfill basic needs of the impaired persons. The operating system is trained and results are obtained



Fig 19 : Output image of hand showing for 1 finger

Here in fig 19 shows the output for first finger which says i need water that is information is conveyed

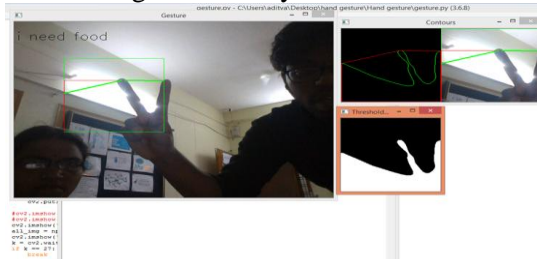


Fig 20: Output image for hand showing 2 fingers

The fig 20 gives information about middle finger which is detected, by eliminating all back ground data tells i need food clearly.

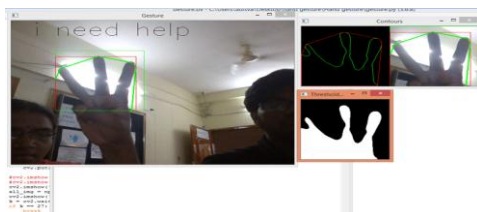


Fig 21: Output image for hand showing 3fingers

In fig 21 shows 3fingers which detects and conveys that i need help, which is basic for disable persons if they are in danger.

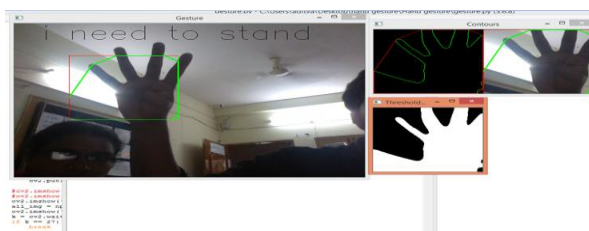


Fig 22 : Output image for hand showing 4 fingers

The fig 22 shows i need to stand for the people who are hospitalised is helpful and conveys information for the staff.

For the data to be tested shows the accurate results we have stored images in database as shown below in fig 23. The tested data is showing in table below in fig 23 for person to person and related graphs are shown. The gesture for each word differs from one to other individuals. The success rate is calculated as

$$\text{Success/Failure} = \text{Recognised} \sim \text{Trained gesture}$$

Based on the above formula the gesture of the persons are compared with other persons in training sets and experimental setups. So as per these, the table is prepared for some words and accordingly pictorised.

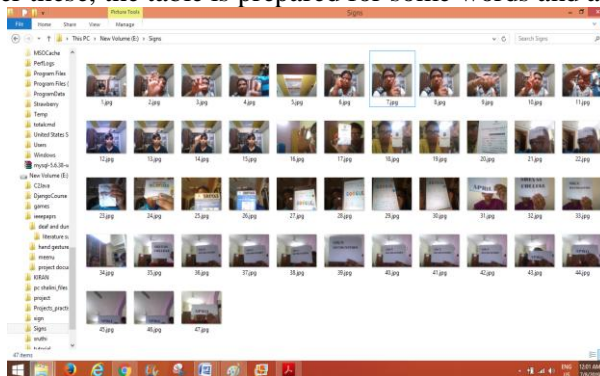


Fig 23: Database of images

Table 1 : success (S) or failure (F) rates of different gestures on different persons

word	person1	person2	person3	person4	person 5	person 6
Hello	S	S	S	S	S	S
Bye	S	S	S	S	S	F
End	S	F	S	F	F	S
Morning	S	S	F	S	F	S
Evening	S	S	S	F	S	S
Sign	S	F	S	S	F	F
Come	S	F	F	F	S	F
Umbrella	F	F	S	F	S	S
Picture	S	S	S	S	F	S

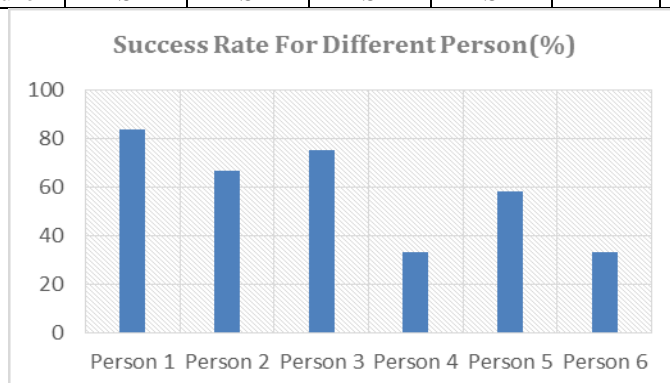


Fig 24: Success rate for different person

For the persons the graph differs in increase and decrease manner ,in which person1 shows all gestures correctly according to the pc but person 4 failed in showing gesture to pc so there is less success rate as represented in fig 24.

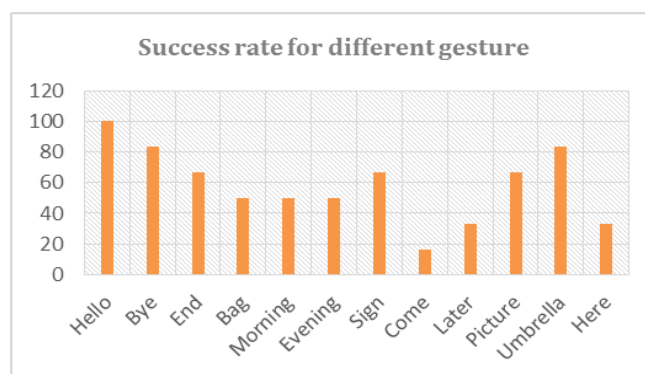


Fig 25:Success rate for different gesture

The words are also tested a per training set and when presented in front of pc the success rate is calculated in which Hello word shows highest rate and whereas come word shows less rate as per fig 25.

CONCLUSION

Gesture based communication Recognition System has been progressed from ordering least complex static indications and letters in order, to the device which can effectively secure powerful activities that is accessible in constant groupings of photos. Analyst as of late is giving more prominent consideration to make a monstrous vocabulary for sign language prominence structures. Numerous specialists are developing their Sign Language Recognition System by methods for the use of little vocabulary and independent database. Huge database build for Sign Language Recognition System keeps on being never again to be had for some of the U.S.A. That worried in creating Sign Language Recognition System. Particularly the Camera-based certainties, which give the shading stream and profundity move video. The sort technique for distinguishing the sign language is in like manner different from analysts. Utilizing their own one of a kind considerations and hindrances for the Sign Language Recognition System, the assessment of procedure to some other strategy stays abstract. Reasonable and direct difference between strategies are compelled in view of the variation of gesture based communication in selective worldwide areas and the qualification in issue set through every analyst. Variety of gesture based communication in the greater part of the U.S.A depends on their syntax and their approach to introduce each word, for example, giving the language through word or by methods for sentence.

FUTURE SCOPE

Gesture based communication Recognition is assuming an overwhelming job in the present situation to decrease gap between the ordinary and debilitated individuals. In future , this division methods utilized for motion acknowledgment framework and furthermore create total communication via gestures acknowledgment framework by utilizing diverse body parts just as outward appearances. The division proficiency of the proposed framework is particularly encouraging for future research in gesture based communication acknowledgment.

REFERENCES

- [1] Dr.ketki prashanth kshirsagar and Dr.Rajeshree Amitkumar Shinde,"Comparing Techniques of Segmenting Hand Region", Proceedings of the 2nd International Conference on Communication and Electronics Systems (ICCES 2017).
- [2] SubhankarChattoraj and Karan Vishwakarmaet ,"Assistive System for Physically Disabled People using Gesture Recognition", In 2017 IEEE 2nd International Conference on Signal and Image Processing.
- [3] Jia Uddin, Fahmid Nasif Arko, Nujhat Tabassum, Taposhi Rabeya Trisha , Fariha Ahmed,"Bangla Sign Language Interpretation using Bag of Features and Support Vector Machine", In 2017 3rd International Conference on Electrical Information and Communication Technology (EICT).

- [4] Rania A. Elsayed, Mahmoud I. Abdalla, Mohammed S. Sayed, "Hybrid Method based on Multi-Feature Descriptor for Static Sign Language Recognition", In the 8th IEEE International Conference on Intelligent Computing and Information Systems (ICICIS 2017).
- [5] Miss. Juhi Ekbote and Mrs. Mahasweta Joshi, "Indian Sign Language Recognition Using ANN and SVM Classifiers", In 2017 International Conference on Innovations in information Embedded and Communication Systems (ICIIECS 2017).
- [6] Naresh Kumar, "Motion Trajectory based human face and hands tracking for sign language recognition", In 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON).
- [7] Mahesh M, Arvind Jayaprakash, Geetha M, "Sign Language Translator for Mobile Platforms", In IEEE 2017 conference.
- [8] Mateen Ahmed, Mujtaba Idrees, Zain ul Abideen, Rafia Mumtaz, Sana Khalique, "Deaf Talk Using 3D Animated Sign Language", In SAI Computing Conference 2016.
- [9] Muhammad Yaqoob Javed, Muhammad Majid Gulzar, Syed Tahir Hussain Rizvi, M Junaid Asif, Zaineb Iqbals, "Implementation of Image processing based digital Dactology converser for Deaf-mute persons", In 2016 IEEE conference.
- [10] Tejas Dharams, Rituparna Jawahar, Kavi Mahesh and Gowri Srinivasa, "Stringing subtitles in sign language", In 2016 IEEE 8th International Conference on Technology for Education.
- [11] Sugunya.R and Dr.T.Meeradevi, "Design of a Communication Aid for Physically Challenged people", In IEEE sponsored 2nd international conference on electronics and communication system (ICECS 2015).
- [12] Prof. Prashant G. Ahire, Kshitija B. Tilekar, Tejaswini A. Jawake, Pramod B. Warale, "Two Way Communicator Between Deaf and Dumb People And Normal People", In 2015 International Conference on Computing Communication Control and Automation.
- [13] Kian Ming Lim, Kok Seang Tan, Alan W. C. Tan, Shing Chiang Tan, Chin Poo Lee, Siti Fatimah Abdul Razak, "Finger Spelling Recognition using Neural Network", In 2015 IEEE Student Conference on Research and Development (SCoReD).
- [14] Sandeep Kaur and Maninder Singh, "Indian Sign Language Animation Generation System", In 1st International Conference on Next Generation Computing Technologies (NGCT 2015).
- [15] Shreyashi Narayna Shawanth and M. S. Kumbhar (2014), "Real Time Sign Recognition using PCA", In IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT).
- [16] C.Jayalaxmi, Dr.Krishnamurthi.V, Dr.A.Revathy (2011), "Transcribing Deaf and Hard of Hearing speech using Hidden Markov Model", In Proceedings of 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN).
- [17] M.K. Bhuyan, Chaitanya Narra and Darsha Sharath Chandra, "Hand Key Gesture using Key Frame Extraction", In 2011 International Conference on Image Information Processing (ICIIP).
- [18] Jian Zaho, Wang Lirong, Zhang Chao, Shi Lijuan, Yin Jia, "Pronouncing Rehabilitation of Hearing-impaired Children based on Chinese 3D Visual-speech Database", In 2010 Fifth International Conference on Frontier of Computer Science and Technology.
- [19] Jiamin Liu, Shuo Chen, Hongxing Sun, Yongxu Qin, Xibo Wang, "Real Time Tracking Method by Using Color Markers", In 2013 International Conference on Virtual Reality and Visualization.

- [20] Akshay Dekate, AnamKamal, Surekha K.S., "Magic Glove-wireless hand gesture hardware controller", In 2014 International Conference on Electronics and Communication Systems (ICECS).
- [21] Remimol.A.M, Sekar.K, "A Method of DWT with Bicubic Interpolation for Image Scaling", In 2014 International Journal of Computer Science Engineering (IJCSSE).
- [22] C. Zhang, Y. Tian, Matt Huenerfauth, "Multi-modality American Sign Language recognition", In Proc. 2016 IEEE International Conference on Image Processing.
- [23] Cheryl M. Lang, Susan Lane-Outlaw, William E. Lange, Dyan L. Sherwood, "American Sign Language/English Bilingual Model:A Longitudinal Study of Academic Growth", In 2013 Journal of Deaf Studies and Deaf .
- [24] Siddharth S. Rautaray, Anupam Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey", In 2015 Artificial Intelligence Review (AIR), Springer .
- [25] Sandeep Kumar, Sukhwinder Singh and Jagdish Kumar "Live Detection Of Face Using Machine Learning with Multi-Feature Method" in Wireless Personal Communication Springer Journal (SCI) DOI: 10.1007/s11277-018-5913-0.
- [26] Sandeep Kumar, Sukhwinder Singh and Jagdish Kumar "Automatic Live Facial Expression Detection Using Genetic Algorithm with Haar Wavelet Features and SVM" in Wireless Personal Communication Springer Journal (SCI) DOI: 10.1007/s11277-018-5923-y.
- [27] Sandeep Kumar, Sanjana Mathew, Navya Anumula, K. Shravya Chandra, "Portable Camera Based Assistive Device for Real-Time Text Recognition on Various Products and Speech Using Android for Blind People" in 8th International Conference on Innovation in Electronics and Communication Engineering (ICIECE-2019).