DIVYANG – A SIGN LANGUAGE RECOGNITION AND CONVERSION SYSTEM

Priyanka
Student, Dept. of Computer Science
ITM Universe Vadodara
Vadodara, India

Krusha Sheth
Student, Dept. of Computer Science
ITM Universe Vadodara
Vadodara, India

Jammy Tandel
Student, Dept. of Computer Science
ITM Universe, Vadodara
Vadodara, India

Avneet Saluja
Asst. Professor, Dept. of Computer Science
ITM Universe, Vadodara
Vadodara, India

Abstract: In today's world the population of approximately 70 million is of deaf and dumb people. Generally deaf and dumb people use sign language for communication i.e. movements of hand to convey their message. One of the important problems that our society faces is that people with disabilities are finding it hard to cope up with the fast growing technology. The access to communication technologies has become essential for these people. Deaf and dumb people always find difficulties in expressing their emotion/ message/thoughts. And some of these difficulties can be easily overcome by Interactive Communication System for Deaf and Dumb people. The system will be software which takes the input of motion of sign language from the user and convert it into meaningful sentences which will be displayed on the screen as output.

Keywords: Interactive Communication, Motion detection system, Sign language, Deaf and Dumb Communication.

I. INTRODUCTION

Normal people have abilities to communicate with each other verbally without any difficulties, but they cannot communicate with people suffering from hearing and vocal problems. These people are also known as Deaf-Mute People. This inability makes them fall apart from being a part of the normal people group. As a result, these people suffer a lot mentally and emotionally as they cannot fully present themselves. However, it is possible to overcome such difficulty and make Deaf-Dumb people able to easily communicate with other people. Different types of systems are present for this particular purpose. Although, up to date technologies are used nowadays but there remain limitations and disadvantages. The first system is Sign Language Converter and this system is based on image processing using an advanced camera, sensor and a computer connected for processing. The second system is Image Processing Based Language Converter system using a computer webcam and a computer to convert RGB images to binary. Our proposed system supports a two-way communication between deaf-mute and normal people.

II. LITERATURE REVIEW

A. Vision-Based Approach

Vision-based approach system requires only the camera to capture the input for human interaction. In this approach, processing is carried out on a pre-stored image. Arpita Ray Sarkar and G Sanyal developed a hand gesture recognition system which consists of models like data acquisition, gesture modeling, feature extraction and hand recognition. In this paper, it is stated that the image is captured using input device which is further used in data acquisition. Then gesture modeling is done which consist of hand segmentation and thresholding. And the features are extracted in order to recognize the hand. In this paper, feature extraction uses hidden Markov model. These systems are simple but face a lot of challenges like a complex background, lighting variation and skin color. [1] [2] [3]

B. Data glove approach

Data glove or hand glove use sensor devices for taking the input which is capturing hand motion The hand glove provides exact coordinates of palm and finger's location, orientation and hand configuration. Five flex sensors are used on each finger to measure the x, y and z position. These coordinates are used for recognition of the gestures and their related letters or alphabets. This data is sent to speech synthesizer to form the speaker so that it can speak that individual word. In another paper by Piyush Kumar, Jyoti Verma and Shitala Prasad use VHand 2.0(5DGT), AD converter, Bluetooth protocol, Mapping Function for human-computer interaction. However, data glove approach demands the user be connected to the computer physically. This makes the interaction between user and computer difficult, hence obstacle the ease of use. [4][5][6]

III. PROPOSED METHOD

In our approach, the user will be using a color band (red, green or blue), so that the system can get an accurate pattern for processing. This paper consists of the detailed concepts of image processing, human-computer interface and

conversion of gesture to text. The proposed system consists of following units, namely:

- Image Acquisition
- Hand Cropping
- Preprocessing
- Feature Extraction
- Classification
- Conversion To Text

1. Image Acquisition

The image acquisition is a unit which deals with the first time interaction process of the user. In this step, a Graphical User Interface (GUI) is prepared which displays the video stream on the screen. From the GUI the count button is clicked which captures an image of the gesture on the screen. The main problem associated with the image captured is that it includes the whole body as well as other unwanted objects. The GUI based front end of the Interactive Communication System is demonstrated in Figure 1.

Figure 1: Image Acquisition

2. Hand Cropping

Once the image acquisition process is completed, the image is processed further for cropping of the hand. This unit detects the actual human hand from other unwanted things which have been captured in the scene. When the portion of the hand is separated from the whole image, the hand is cropped out. For the separation of the hand and cropping it out, a certain threshold is used. Overlaid onto the image with human skin pixels are marked in blue color so that the gesture can be easily identified.
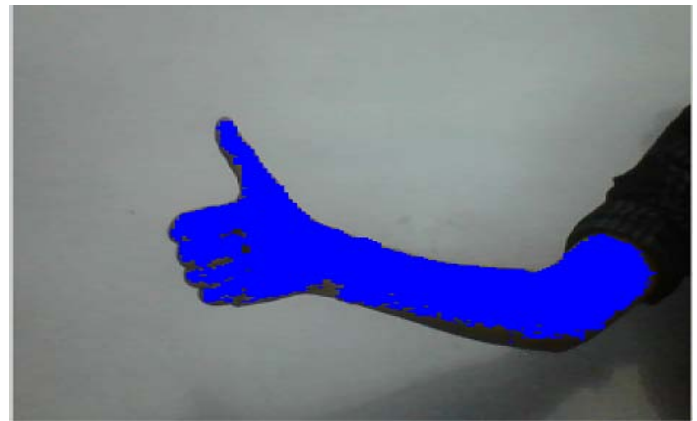

Figure 2: Hand Cropping

3. Preprocessing

3.1 RGB to Grayscale and Binarize.

RGB (Red, Green, and Blue) is a system of colors. These are the 3 primary colors available which are added in different quantities to give different types of other colors. The humans have vision ability to distinguish between many colors, their intensities, and their shades. But gray is the only color available where humans can only differentiate around 100 shades of gray. Therefore it is evident that the images that are colored contains more information, and is to be converted into grayscale by using grayscale filtering. In Digital Image Processing domain, there are different types of filters.

The conversion of grayscale image to binary image is termed as Binarization. In a gray level image, the level varies from 0 to 255, whereas in binary images it is 0 & 1. The 0 and 1 in the binary image represent two different colors i.e. black and white. This will help in decreasing the time consumed during feature extraction process.
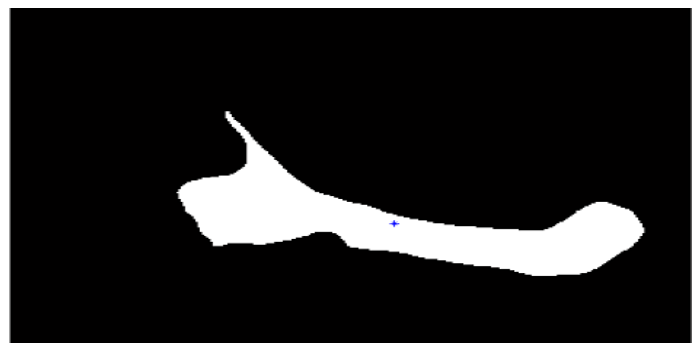

Figure 3: Binarization Process

3.2 Removal of noise from the image.

Noise is the undesired variation or changes in the brightness, as well as the color of an image. The removal of noise from an image is very much necessary, as it will affect the results. If the feature extraction process takes place without the noise removal process, then it will result in incorrect classification. Therefore in order to prevent this misleading, the image undergoes the Noise Removal process. This step helps in improving the result.
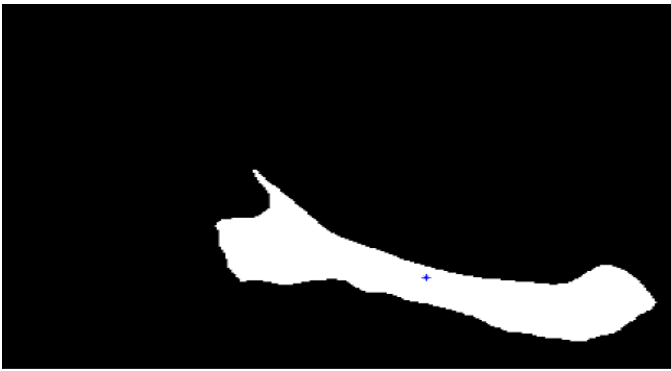
Figure 4: Removal of noise from the image

4. Feature Extraction

A feature is a small information about the image or the minor and important details about the image. These details can be anything like objects, edges, etc in the image. To find the movement information, the input gesture is assumed to be non-stationary or moving. When the object moves in the spatial-time space, an image sequence is generated and is tracked by the motion detector by examining the local gray scale level. There are various algorithms used for feature extractions such as Zernike moments and Fourier descriptors. The feature extraction unit includes thresholding, edge detection, motion and region identification.

5. Classification

The classification unit includes the recognition learning phase. In this recognition learning phase, various steps are followed in order to successfully extract the feature and recognize the exact meaning of the image. First, the features of the test image are calculated, and these features are compared with the training feature set. For classification, an algorithm is used, named as K-nearest neighbor (KNN) which uses neighbors to calculate the distance. After calculating the distance the value is compared with the threshold; if it passes the threshold it is classified or otherwise, it is identified as a new gesture.

6. Conversion to Text

Once the gesture is processed through the mentioned steps, the class of the gesture which was given at runtime is obtained. In this unit, all types of texts are available for every particular gesture. The classification of each type of gesture is predefined in the system, which provides an accurate result for the input. Then after the meaningful sentence is displayed on the screen by matching with the predefined or stored gesture.

IV. RESULT

In the proposed system, the pre-processing modules are capable of converting the images into binary form i.e. in a black and white image. The black part of the image will determine the unwanted section of the image whereas the white part will be the important section of the image. All

these gesture/ images are recognized and match with the database to identify the correct meaning of the input. The accuracy of the proposed system is above 95% and is capable of converting the sign language gestures into meaningful sentences. The system uses Sign Language gesture conversion in English language only. It can also be created for other languages as well. The output images of general sentences displayed as well as the conversion of gesture to a meaningful text are as shown below:
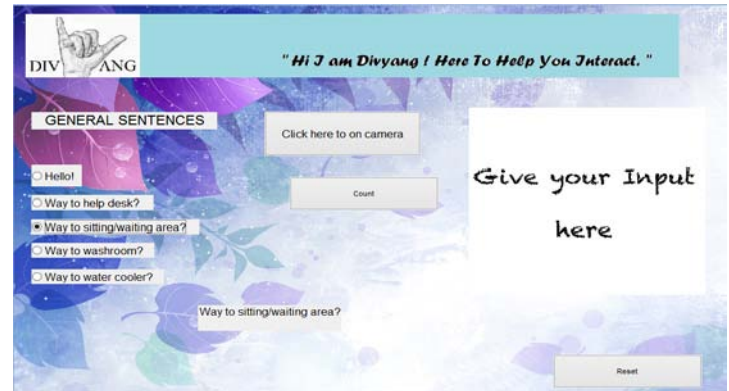
a.) General Sentence Display

Figure 5: General sentence displayed on the screen
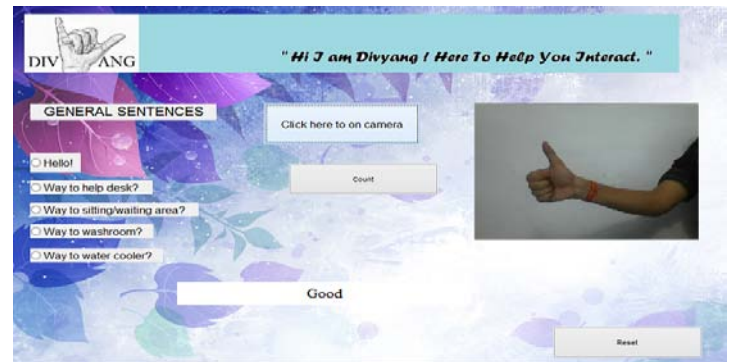
b.) Gesture to Text Display

Figure 6: Gesture to meaningful text displayed on the screen

V. CONCLUSION

The system supports a friendly and portable two-way comprehensible communication between deaf-dumb and normal people. This Interactive Communication System converts the sign language gesture inputs into meaningful texts. The system design takes into consideration that it can be easily installed and used in different domains and used by deaf-mute people. Besides, the system contains an enormous sign language data stored in the MATLAB system which makes the system reliable to be used. The output can be easily displayed on any screen as desired by using a microcontroller. The processing speed of the system is vastly fast as compared to current system.

VI. FUTURE SCOPE

1. By integrating our interactive communication system with voice recognition system we can embed it in Robots.

2. Different languages like Hindi, State languages, etc can be used for displaying the output.

3. For achieving higher accuracy, a better and more accurate image acquisition device can be used.

VII. ACKNOWLEDGEMENT

REFERENCES

[1] Chen Yong Qiang, HUANG YingShu, Xia Chenjie "Static hand gesture recognition algorithm based on finger angle characteristics", Proceedings of the 33rd Chinese Control Conference, pp 8239 - 8242, 2014.

[2] Joyeeta Singha, and Karen Das, "Recognition of Indian Sign Language in Live video", International Journal of Computer Applications", Vol.70, No.19, pp 17-22, May 2013.

[3] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat "Vision-Based Hand Gesture Recognition", World Academy of Science, Engineering and Technology 49, pp. 972-977, 2009.

[4] G. R. S. Murthy & R. S. Jadon, "A Review of Vision-Based Hand Gestures Recognition", International Journal of www.ijcat.com 17 Information Technology and Knowledge Management, vol. 2(2), pp. 405-410, 2009.

[5] Shitala Prasad, Piyush Kumar, Kumari Priyanka, "A Wireless Dynamic Gesture User Interface for HCI Using Hand Data Glove", Contemporary Computing (IC3), 2014 Seventh International Conference, pp. 1617-1621,2014.

[6] Piyush Kumar, Jyoti Verma" Hand Data Glove: A Wearable Real-Time Device for Human-Computer Interaction" International Journal of Advanced Science and Technology Vol. 43, pp-15-26, June 2012.