Kinect Sensor based Gesture Recognition for Dumb People with Voice Extraction

Sapana Jindam[1] Babita Kolekar[2] Archana Pawar[3] Lalita Palange[4]
[1,2,3]U.G. Student [4]Associate Professor
[1,2,3,4]Department of Electronics and Telecommunication Engineering
[1,2,3,4]SVERI's College of Engineering, Pandharpur, Solapur University, Solapur, Maharashtra, India

Abstract— Nowadays, we mostly hear the word new Technology. This word advancing the old technologies along with some extra features. This growing towards new discoveries and inventions in the field of science and technology but unfortunately, there are very fewer innovations for physically challenged people. Who face difficulties in communicating with normal people. As Disable people use sign language as their Prime medium for communication. Mostly, sign languages are not understood by the common people. Research says that many studies have been done to eliminate such kind of communication barrier. And so, many research works have been proposed to remove this communication gap between normal people and physically impaired people. So, In this Paper, we are introducing one idea which will remove the complexity in earlier proposed methods such as, Microcontroller's, Aurdino, flex sensor based hand gloves. We are advancing this process in our study by using a Kinect sensor. Kinect is a motion sensing input devices produced by Microsoft. Many applications out of which motion sensing and 3-D image processing is the main requirement for our work to detect the signs, gestures, and movements of hands or bodies. Our work starts from capturing an image of the body to convert into the skeletal image and from image processing to feature extraction of the detected image hence getting an output along with voice. The experimental results of our Presented algorithm are also very good with an accuracy of 90%.

Keywords: Image Processing, Kinect Sensor, Skeletal Image

I. INTRODUCTION

As we see in our surrounding, many Deaf and Dumb people are unable to share their views to normal people and Sometimes people interpret their messages wrongly either through sign language or through lip reading or lip sync. In recent years, there has been a rapid increase in the number of deaf and dumb victims due to birth defects, accidents and oral diseases. Since deaf and dumb people cannot be able to communicate with a normal person so they have to depend on some sort of visual communication. So, This project is made in such a way to help these especially challenged people & hold equal priority in society.

HGRVC (Hand Gesture Recognition and Voice Conversion) system tracks the hand gestures of the dumb and deaf people in order to maintain a communication channel with the people. The detection of hand gestures can be done using a Kinect sensor. By this, it senses the data and a 3-D image are created. This data is then transferred to Matlab where it is interfaced through the programming along with image processing and feature extraction. Only the image of the hand is cropped from whole body image, the gesture of that hand is then equated with the available image in the database and if they match the speech and text is obtained as output making it easier for the common people to understand it. By this, impaired people will be confident to express their views anywhere and everywhere.

After the gestures are captured the images are transformed from one form to another for the final detection of the sign. The whole process runs from transforming the skeletal image to a segmented image then the segmented image is cropped. The cropped a segmented image is then converted into dots and dashes image.

Below figure shows Kinect sensor assembly. It consists of its all internal parts shown.

Fig. 1: Kinect Sensor

II. RELATED WORK

The detailed literature review has been studied and some of the most sought studies are presented here.

Jadhav at el. [1] Presented "Hand Gesture Recognition to Speech Conversion in Regional Language": This methodology provides a map for developing a Digital wireless glove which is fitted with Flex sensors and accelerometer. This system includes a voice playback IC to give real-time speech output. According to Author, to increase Capability they must have to increase No. of flux sensors and Accelerometer.

Channaiah at el. [2] developed "Hand Gestures Recognition System for Deaf and Dumb and Blind People": They use Web Camera and Microcontroller. In this they Design a code to detect colours and to send the corresponding ids. According to them, Their Project gives the speech as output and must be analysed and converted to textual display on the screen of ALCD.

Sunita at el.[3] developed "Electronic Support System for Deaf and Dumb to Interpret Sign Language of Communication": In this, they used flex sensor and ARM7TDMI. According to the Author, their Project will only recognize Alphabets and numbers.

V. Padmanabhan, M. Sornalatha [6] presented "Hand gesture recognition and voice conversion system for

dumb people": In this Paper, they used a Flex sensor for sensing the action its output is given to microcontroller for matching input with the database. According to them this Project overcomes the necessary Time Difficulties of Dumb People and Improves their manner.

Prajapati at el. [7] Investigated "Hand Gesture Recognition & voice conversion For Deaf and Dumb People": In this Paper, they have used Web Camera and PCA Algorithm. According to Author implementation of this system gives up to 90% accuracy and works successfully in most of the test cases.

Jagadish at el. [8] Presented "Hand Gesture Recognition System for Deaf and Dumb Persons": In this project, they used the methodology which provides a map for developing a digital wireless glove which is fitted with Flex sensors. This system includes a voice playback IC to give the real-time speech output in regional language. According to Author, This paper has the potential of minimizing this communication barrier by working as an automated translator and converting sign language directly into a vocal and textual format.

Kunal at el. [9] investigated "Sign Language Recognition For Deaf And Dumb People Using ANFIS": In this they used International Neural Networks, Hidden Markov Model (HMM) According to author, their system gives maximum efficiency, having low cost, and is an optimal mixture of methods, giving results against complex backgrounds as well, should be preferred.

Jain at el. [4] proposed "Image Processing based speaking system for mute People using hand gesture": In this system, a webcam will capture the hand gesture and perform image processing using principal component analysis algorithm (PCA). The Author says, their design is more compatible and faster responsive when compared to an existing design using PCA.

Pradeep and Tharanyaa [5] presented "An artificial speech system for dumb and blind people": Here they used an accelerometer sensor, flex sensor & ultrasonic sensor. According to the author, this project will mostly helpful for deaf and dumb people.

Dawan at el. [10] developed "A Review on Hand Gesture Recognition for Deaf and Dumb People Using GSM Module": In this they used flux sensor along with gloves and GSM modem for displaying messages. According to the author, their Project requires number of flex sensor for high accuracy.

III. PROPOSED ALGORITHM

A. Minimum Distance Classifier:

The minimum distance classifier is used to classify unknown image data to classes which minimize the distance between the image data and the class in multi-feature space. The distance is defined as an index of similarity so that the minimum distance is identical to the maximum similarity. As the classification is based on the minimum distance calculation, this method is called minimum distance classification procedure. Minimum distance procedure performs efficiently if pattern classes can be represented by a single prototype or by several prototypes around which the patterns form clusters. The simplest type of minimum distance classifier is the one wherein the patterns of all classes are very close to each other and each class can be represented by a single prototype.

B. Image Matching Process:

Feature extraction a type of dimensional reduction that efficiently represent interesting part of an image as a compact feature vector and Detecting an object (left) in a cluttered scene (right) using the combination of feature detection, feature extraction and matching.

We have proposed an algorithm which follows the following steps:

1) In this first step, we have to perform gesture in front of Kinect sensor so that it will capture the image and it locates the joints of the body by pointing it out and hence we get a skeletal image.
2) Then the segmented image of the body is formed from the skeletal image. The area of the hands where the signs are captured are cropped out of the whole segmented image. The length of the dash is 4 units and the spacing between the two lines of dashes is also 4 units.
3) Through observation, we found that wherever the length of the dashes is greater than 4 units it resembles the image of the cropped hand.
4) In our proposed algorithm, we have taken the concept of loops, it is used to detect the black points that are the space between the dashes. This detection of black points determines the position of dashes by successive subtraction of points in the iterations which is going on and on.
5) The basic algorithm behind this work is that after successive subtraction of points if the value is equal to 4 then there is no data and if the value is greater than 4 then there is the actual image of the cropped hand.
6) Now, here arises a problem that how we detect the location of fingers. So the algorithm behind this is based on the formation of matrices of the black points detected earlier. In an iteration, the matrix coordinates of the finger are four times the number of lines of dashes.
7) To highlight the fingers, we plotted star point of the same coordinates as of the fingers and for more precise feature extraction the image processing is followed by filtration of the image that means the star points plotted on the figures go through following conditions. If they fall on a straight line either horizontal or vertical. If they fall on a constant slope. If they fall within 4-unit coordinate difference.
8) Then it eliminates the identical points which we call as garbage point. Now we get the filtered image but the process of feature extraction continues for identification of fingers that whether it is an index finger, middle finger, ring finger, little finger or thumb.

Flow chart of the above algorithm is shown below:

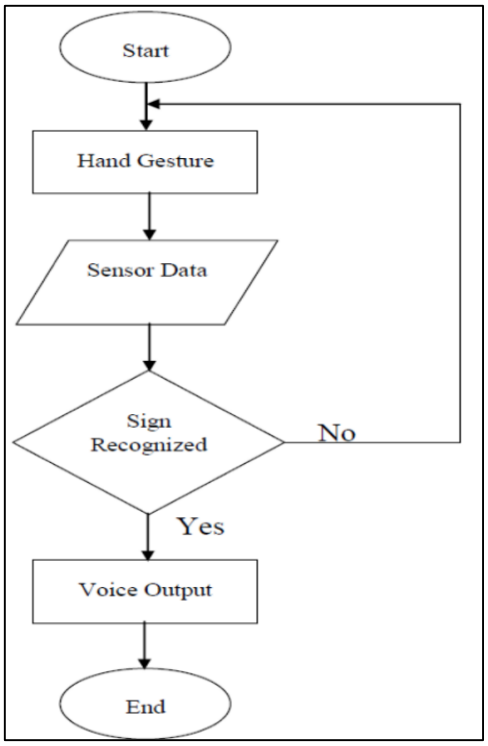

Fig. 2: Flow chart

IV. DATABASE

The Database can be created efficiently by a user which is used for training the system, the only limitation while creating database is that background should white and clear. The user has to specify the number of samples; a number of samples basically means a number of images in the database. It is required to make a proper database of gestures so that the image captured using this system can be compared. For making the database, we would have to capture each gesture from more than 2 angles so that the accuracy of the system will increase significantly. We get the database from renowned resources so that we can validate our results with the standard result already available, this will help us to validate and verify our results with existing results.

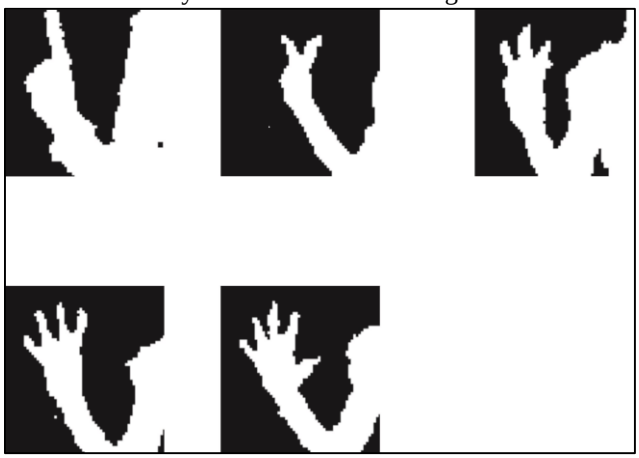

Fig. 3: Database images

V. SIMULATION RESULTS

Finally, after the detection of the whole image of fingers or we can say that a complete hand ANDing operation continues in Matlab for the final output. The detected image of the sign is searched in the database for its meaning and as and when the match is found search is complete and we get the final output along with the image of the meaning of sign and its voice. We can take the example as, if number 4 is to be detected by the Kinect sensor then the person gestures 4 using his hands. The Kinect captures the skeletal image of the body as shown in Fig.4

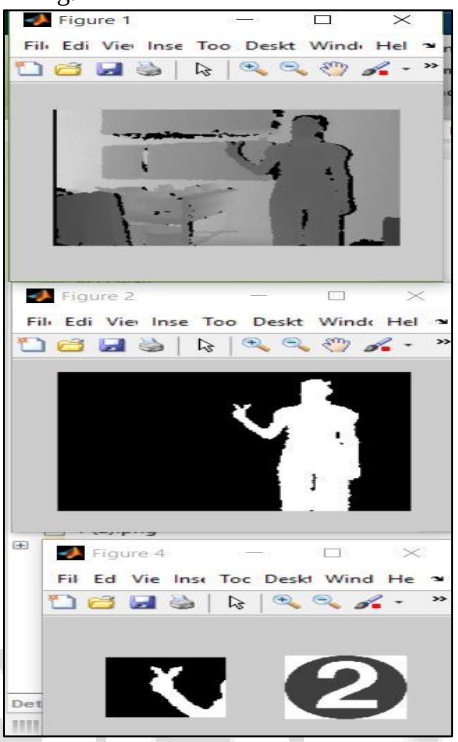

Fig. 4: Output for sign 2.

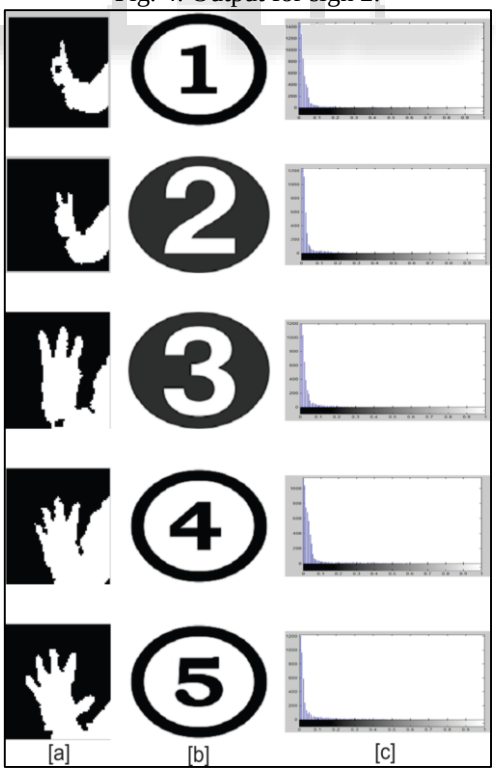

Fig. 5: Output for 1 to 5 signs: (a) hand input given. (b) Output voice with No. (c) Histogram of sign.

VI. CONCLUSION

After referring all the earlier studies, our work has given more precise output. Since programming is done for the detection of the left hand and the coordinates are taken accordingly. This our Idea is removing complex circuitry and our algorithm gives all the relevant information about the coordinates of each and every finger detected. This system is easily understandable, flexible and user-friendly as the user can be of any age, gender, size or color, the results will be the same. But the intensity of light and distance of the body from the sensor affects the efficiency. For higher efficiency with the devices, it is suggested that keep the Kinect sensor at a height of about 65 cm from the ground and the person should stand distanced at 90 cm from Kinect.

REFERENCES

[1] B.D. Jadhav, N. Munot, M. Hambarde, and J. Ashtikar, "Hand gesture recognition to speech Conversion in regional language", International Journal of Computer Science and Network, vol. 04, pp. 161–166, 2015.
[2] C. Chandana, K. Nikhita, P. Nikiitha, N. K. Bhavani, and J. Sudeep, "Hand gestures recognition system for deaf and dumb and blind people", International Journal of Innovative Research in Computer and Communication Engineering, vol.05,pp. 0058–10062, 2017.
[3] D. M. Sunita V. Matiwade1, "Electronic support system for deaf and dumb to interpret sign language of communication", International Journal of innovative Research in Science and Engineering and Technology, vol. 05, pp. 8683–8689, 2016.
[4] A. Jain, L. Jain, I. Sharma, and A. Chauhan, "Image processing based speaking system for mute people using hand gestures", International Journal of Engineering Sciences & Research Technology, vol. 03, pp. 368–374, 2018.
[5] U. Pradeepa and J. P. Tharanyaa, "An artificial speech system for dumb and blind people", International Journal of Advanced Research in Electronics and communication Engineering (IJARECE), vol. 05, pp. 476–479, 2016.
[6] V. Padmanabhan and M. Sornalatha, "Hand gesture recognition and voice conversion system for dumb people", International Journal of Scientific Engineering Research, vol. 05, pp. 427–431, 2014.
[7] R. Prajapati, V. Pandey, N. Jamindar, N. Yadav, and N. Phadnis, "Hand gesture recognition and voice conversion for deaf and dumb", International Research Journal of Engineering and Technology (IRJET), vol. 05, pp. 1373–1376, 2018.
[8] R. Jagadish, R.Gayathri, R. Mohanapriya, R. Kalaivani, and S. Keerthana, "Hand gesture recognition system for deaf and dumb persons", Indo-Iranian Journal of Scientific Research (IIJSR), vol. 02, pp. 139–146, 2018.
[9] M. K. A. Wankhade and P. G. N. Zade, "Sign language recognition for deaf and dumb people using anfis" , International Journal of Science, Engineering and Technology Research (IJSETR), vol. 03, pp. 1206–1210, 2014.

