

# Hand Gesture Recognition for Deaf and Dumb

Parshwa P. Patil<sup>1</sup>, Maithili J. Phatak<sup>1</sup>, Suharsh S. Kale<sup>1</sup>, Premjeet N. Patil<sup>1</sup>, Pranav S. Harole<sup>1</sup>,  
Prof. Dr. C. S. Shinde<sup>2</sup>

<sup>1</sup>Student, Dr J. J. Magdum College of Engineering, Jaysingpur, Maharashtra, India

<sup>2</sup>Asst. Professor, Dept. of Computer Science and Engineering, Dr J. J. Magdum College of Engineering, Jaysingpur, Maharashtra, India

\*\*\*

**Abstract** - The vision-based technology of hand gesture recognition is an important part of human-computer interaction (HCI). Owing to the rapid development of hardware and software, new types of HCI methods have been required. Technologies such as speech recognition and gesture recognition receive great attention in the field of HCI. Deaf and dumb individuals lack in proper communication with normal people and find it difficult to properly express themselves. Thus, they are subjected to face many issues in this regard. The sign language is very popular among them and they use it to express themselves. The deaf and dumb make use of sign language to communicate which is difficult to interpret by the individuals who are not well-aware of it. There is a need of building up a device that can interpret the gestures into text and speech. This will be a great step to make the communication possible between the deaf and dumb individuals and the general public. We are providing an interface for deaf and dumb people who cannot speak to interact with other people in their surroundings.

**Key Words:** sign language; deaf; dumb; hand gesture; image recognition; feature matching; feature extraction; gesture recognition

## 1. INTRODUCTION

Gesture is a symbol of physical behaviour or emotional expression. It includes body gesture and hand gesture. It falls into two categories: static gesture and dynamic gesture. For the former, the posture of the body or the gesture of the hand denotes a sign. For the latter, the movement of the body or the hand conveys some messages. Gesture recognition determines the user intent through the recognition of the gesture or movement of the body or body parts. In the past decades, many researchers have striven to improve the hand gesture recognition technology. Hand gesture recognition has great value in many applications such as sign language recognition, augmented or virtual reality, sign language interpreters for the disabled, and robot control. The importance of gesture recognition has increased at a very fast pace owing the new generation of gesture interface technology. Sign language is mainly a language for the differently-abled deaf-mute community who can't use spoken languages to communicate with others. Although they can see, the use of hand signs to interact becomes inconvenient to express one's feelings readily if a common hand sign language is not followed. A standard sign language has a defined set of signs and their meanings,

making them easy to understand. The sign language too uses different gestures for communication mostly the hand. Sign languages are distinct than spoken languages. The use of hand gestures for interaction either with humans or with machines is comparatively higher than other body gestures like head and eyes due to the fact that hands send more clearer signals and the gestures can be made spontaneously. Nowadays, deaf and dumb people are becoming more and more outgoing, and unlike olden days, do not depend on anybody for communication. So, for such people it is important that the general public around them must also be able to understand what they want to tell them using sign language. For such people, we build the hand gesture recognition program so that even if anyone does not know the meanings of the signs, they can use this program to recognise the sign language. Hence, the deaf and dumb people will have no barrier in their communication.

## 2. LITERATURE SURVEY

Nowadays, every electronic device around us can be controlled without any need to touch it, due to the surge of voice and gesture controls. Voice controls recognize the words spoken by a person. While in gesture controls the device is controlled by moving our hand in a specific pattern for the device to perform a particular task. A lot of research has been carried out for designing a fast and efficient hand gesture recognition system. There are many experiments for recognizing hand gestures, and such programs are also being used as parts of other programs such as controlling a device using hand gestures, etc. There are many different ways to recognize a hand gesture from an image.

[1] The deaf and dumb make use of sign language to communicate which is difficult to interpret by the individuals who are not well-aware of it. Thus, there is a need of building up a device that can interpret the gestures into text and speech. This will be a great step to make the communication possible between the deaf and dumb individuals and the general public.

[2] Sign language recognition involves hand gesture classification. A method for static hand gesture classification has been proposed for American Sign Language (ASL) finger spelling alphabets and digits. The system uses skin color-based segmentation requiring little post processing. The averages of central moments of order 2 to 9 have been extracted as the features that characterize

the hand gestures. The neural network classifier has been employed for recognition and it produced decent classification results of 73.68% with a small feature vector size containing 8 features.

[3] The work flow of hand gesture recognition is described as first, the hand region is detected from the original images from the input devices. Then, some kinds of features are extracted to describe hand gestures. Last, the recognition of hand gestures is accomplished by measuring the similarity of the feature data. The input devices providing the original image information includes normal camera, stereo camera, and ToF (time of flight) camera. The stereo camera and ToF camera additionally provide the depth information so it is easy to segment the hand region from the background in terms of the depth map.

[4] To recognize different hand gestures and achieve efficient classification to understand static and dynamic hand movements used for communications. Static and dynamic hand movements are first captured using gesture recognition devices including Kinect device, hand movement sensors, connecting electrodes, and accelerometers. These gestures are processed using hand gesture recognition algorithms such as multivariate fuzzy decisions tree, Hidden Markov Models (HMM), dynamic time warping framework, latent regression forest, support vector machine, and surface electromyogram. These captured gestures are processed for occlusion and fingers close interaction for identification of right gesture capture devices with proper illumination conditions. These captured gestures are processed for occlusion and fingers close interactions for identification of right gesture and to classify the gesture and ignore the intermittent gestures. Real-time hand gestures recognition needs robust algorithms like HMM to detect only the intended gesture. Classified gestures are then compared for the effectiveness with training and tested standard datasets like sign language alphabets and KTH datasets. Hand gesture recognition plays a very important role in some of the applications such as sign language recognition, robotics, television control, rehabilitation, and music orchestration.

[5] Gesture is the most primitive technique for conversation amongst human being. Today in the time of current day innovation gesture recognition affects the world differently, from the physically challenged individuals to robotic manage to virtual fact conditions. Human hand gestures provide the natural and advantageous approach of non-verbal conversation with the computer interface. Hand gestures are the considerable body moves that are actions of hands, arms or fingers. Hand gesture recognizable proof levels from the static gesture with the complicated foundation or dynamic gestures that express the human feeling and communicate with computer or humans. The hand is specifically use as the contribution to the machine, for the verbal exchange reason for gesture identification there is no need of an intermediate medium. In this paper, a deep convolution neural network is proposed to immediately classify hand

gestures in pictures without any segmentation or detection stage that could discard the irrelevant not-hand area.

### 3. PROPOSED FRAMEWORK

We are going to recognize hand gestures from a video sequence. To recognize these gestures from a live video sequence, we first need to take out the hand region alone removing all the unwanted portions in the video sequence. After segmenting the hand region, we then count the fingers shown in the video sequence to instruct a program based on the finger count. Thus, the entire problem could be solved using 5 simple steps: First, we need to find and segment the hand region from the video sequence. Next, count the number of fingers, area of palm from the segmented hand region in the video sequence. Then we will match the segmented part to the available dataset. Then, select the most accurate data from the dataset and assigned a weight for next comparison, for faster selection of data. Finally, we will convert the obtained data into the text and display the text. Required formula [3] to find out defect point is given below:

$$\cos((a^2 + b^2 - c^2)/2bc) \times 180 + \pi$$

#### 3.1 Relevance of the Work

As deaf and dumb people cannot speak like normal people, they use hand sign language to communicate with other people like themselves. But when it comes to communicating with normal people, it proves to be difficult, as most normal people do not understand the sign language used by them. So, we are developing this program to make it easier for the deaf and dumb to communicate with the people who do not understand their sign language. Our program will detect and translate the hand signs made by the user into text and display it on the screen for the normal people to read. Thus, our project will make it easier for deaf and dumb to communicate with the normal people.[1][2][5]

#### 3.2 Proposed Work

In this project, we propose the conversion of hand gesture into text for deaf and dumb people. The main theme of our project is to recognize the hand gesture, detect the gesture, and show the output in the form of text. The end user has to perform hand gestures in front of the camera. As user perform the gestures, our program will detect the gesture and it will translate the gesture into text in parallel [2]. There will be two display on the screen, one will show the hand gesture and the other will show the text represented by that gesture. Our project works as a translator for the deaf and dumb. It reduces many problems such as the need of a human translator. Deaf and dumb people will be able to express themselves through our program. We are going to recognize hand gestures from a video sequence. To recognize these gestures from a live video sequence, we first need to take out the hand region alone removing all the unwanted portions in the video sequence. After segmenting the hand region, we

then count the fingers shown in the video sequence to instruct a program based on the finger count. Thus, the entire problem could be solved using 5 simple steps: First, we need to find and segment the hand region from the video sequence. Next, count the number of fingers, area of palm from the segmented hand region in the video sequence. Then we will match the segmented part to the available dataset. Then, select the most accurate data from the dataset and assigned a weight for next comparison, for faster selection of data. Finally, we will convert the obtained data into the text and display the text. The end user has to perform hand gestures in front of the camera. As user perform the gestures, our program will detect the gesture and it will translate the gesture into text in parallel. There will be two display on the screen, one will show the hand gesture and the other will show the text represented by that gesture. Our project works as a translator for the deaf and dumb. It reduces many problems such as the need of a human translator. Deaf and dumb people will be able to express themselves through our program [1][2][3][4][5].

### 2.3 System Architecture

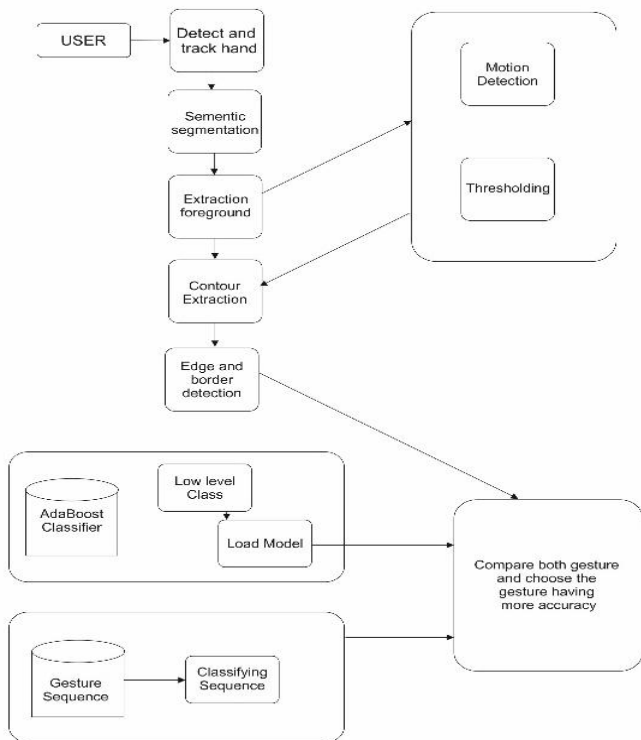


Fig -1: System Architecture

### 3.3 Techniques to be used

#### 3.3.1. AdaBoost Classifier

AdaBoost, short for Adaptive Boosting, is a machine learning meta-algorithm. It can be used in conjunction with many other types of learning algorithms to improve performance. AdaBoost is adaptive in the sense that subsequent weak learners are tweaked in favour of those

instances misclassified by previous classifiers. AdaBoost is sensitive to noisy data and outliers. In some problems it can be less susceptible to overfitting problem than other learning algorithms. AdaBoost training process selects only those features known to improve the predictive power of the model, reducing dimensionality and potentially improving execution time as irrelevant features need not be computed.

#### 3.3.2 Region of Interest

A region of interest (often abbreviated ROI), are samples within a dataset identified for a particular purpose. Region of interest pooling (also known as ROI pooling) is an operation widely used in object detection tasks using convolution neural networks. For example, to detect multiple cars and pedestrians in a single image. Its purpose is to perform max pooling on inputs of non-uniform sizes to obtain fixed size feature maps.

The object detection architecture is broken down in following stages:

**i. Convex Hull:** A convex object is one with no interior angles greater than 180 degrees. a shape that is not convex is called Non-convex. The convex hull is a ubiquitous structure in computational geometry. Even though it is a useful tool in its own right, it is also helpful in constructing other structures like Voronoi diagrams, and in applications like unsupervised image analysis.

**ii. Contour extraction:** Edge detectors typically produce short, disjoint edge segments. These segments are generally of little use until they are aggregated into extended edges. Two main categories of methods are: local methods (extended edges by seeking the most “compatible” candidate edge in a neighborhood), global methods (more computationally expensive-domain knowledge can be incorporated in their cost function).



Fig -2: Contour Extraction

#### 3.3.3 Feature Extraction

In machine learning, pattern recognition and in image processing, feature extraction starts from an initial set of measured data and builds derived values(features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations [1]. When the input data to an algorithm is too large to be

redundant then it can be transformed into a reduced set of features. Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data. Feature extraction involves reducing the number of resources required to describe a large set of data.

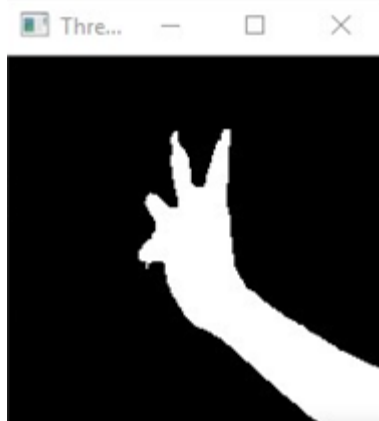
### 3.3.4 Motion Detection

Motion detection is the process of detecting a change in the position of an object relative to its surroundings or a change in the position of an object relative to its surrounding or a change in the surroundings relative to an object. Motion detection can be achieved by either mechanical or electronic methods. Motion detection is usually a software-based monitoring algorithm which, when it detects motions will signal the surveillance camera to begin capturing the event.

### 3.3.5 Thresholding

Thresholding is a method of image segmentation, in general it is used to create binary images. Thresholding is of two types namely, simple thresholding and adaptive thresholding.

Simple Thresholding: In simple thresholding operation the pixels whose values are greater than the specified threshold value, are assigned with a standard value. You can perform simple threshold operation on an image using the method *threshold()* of the *Imgproc* class.



**Fig -3: Thresholding**

## 4. CONCLUSION

The proposed hand recognition system is very useful as it can be used as human-computer interface as well as can help paralyzed people. In future devices can be easily controlled through the hand gesture movement easily. In previous concepts, the sensor techniques were used but, in this system, no any additional sensor is needed. It can be achieved by deep convolutional neural network algorithm. Deep learning deals with heavy data set. The convolutional network is layered algorithm which is mainly divided into

two phases. First one is feature extraction with convolutional neural network and another one is classification phase with fully connected layers. The training of neural network is done for a heavy set of data and goal is achieved.

## REFERENCES

- [1] Anchal Sood, Anju Mishra, "AAWAAZ: Communication System for Deaf and Dumb", Student, Amity School of Engineering and Technology, Amity University, Noida, India aanchalplp@gmail.com.
- [2] Sunanda Biradar, Ashwini M. Tuppad, "A Static Hand Gesture Classification system for American Sign Language Fingerspelling and digits", International Journal of Latest Trends in Engineering and Technology (IJLTET) ISSN: 2278-621X.
- [3] Zhi-hua Chen, Jung-Tae Kim, Jianning Liang, Jing Zhang and Yu-Bo Yuan, "Real-Time Hand Gesture Recognition Using Finger Segmentation", Hindawi Publishing Corporation Scientific World Journal Volume 2014, Article ID 26787.
- [4] Srinivas K, Manoj Kumar Rajagopal, "Study Of Hand Gesture Recognition and Classification", School of Electronics Engineering, VIT University, Chennai, Tamil Nadu, India. Email: manojkumar.r@vit.ac.in.
- [5] Kundan Kumar Dubey, Ajitanshu Jha, Akshay Tiwari, K. Narmatha, "Hand Gesture Movement Recognition System Using Convolution Neural Network Algorithm", International Research Journal of Computer Science (IRJCS) Issue 04, Volume 6 (April 2019) ISSN: 2393-9842.