

Multimedia education platform for hearing impaired students in India with bidirectional translation of regional sign language

Mr. Sandip R. Shetake

Assistant Professor Department of computer Science & Engg. TKIET, Warananagar.

Abstract - We can express our views by talking through our voice with other persons but right now in a world there are Deaf and Dumb persons who can't express their views through voice but they are using sign language to express their views which can be understandable for other deaf and dumb persons but if we want to make the conversation with such kind of impaired persons normal persons can't easily understand what they mean to say, so the proposed work is understand the sign language and create a platform which will capture the signs and convert them in our understandable format also it will help the trainers to teach proper sign language to deaf and dumb persons.

Key Words: Indian Sign Language (ISL), Otsubinary Kohonen

1. INTRODUCTION

Sign languages are natural languages that use different means of expression for communication in everyday life. More particularly, it is the only means of communication for the hearing impaired. Thus, it offers enhancement of communication capabilities among normal beings and provides replacement for speech among deaf and mute people. Because of these, automatic sign language recognition has attracted vision researchers for long. Several research works are going on sign language recognition in order to make the communication between a deaf person and a normal person easy.

Sign languages are well structured languages with a phonology, morphology, syntax and grammar distinctive from spoken languages. The structure of a spoken language makes use of words linearly, i.e., one after the other, whereas a sign language makes use of several parallel body movements in the spatial as well as in temporal space.

1.1 Existing System:-

The existing work related to Sign Language Recognition has been carried out earlier as well but it makes use of some of the costly devices and having some limitations.

a) Kinetic Sensor:- Kinect is Microsoft motion sensor it consist of RGB camera , depth sensor and multi-array microphone. It recognizes facial movement and speech. Depth camera is kinect sensor used to detect ASL alphabet.

Distance adaptive scheme was used for feature extraction. Support vector machine and RF classifier algorithm used for classification purpose. Training of data was done using

ANN network. The accuracy of the system was 90%. Firstly it detects hand movement and then matched with counter model. Second task was to locate multi color glove and detect different color regions. This system has one drawback that is limited accuracy.

b) Data Glove: - This method uses different sensor to detect hand gesture signal. Hand gesture signal is in the form of analog. ADC is used to convert analog signal into digital form. It consists of flex sensor and accelerometer.

Flex sensor is used to detect bend signal. The system consists of flex sensor, accelerometer and tactile sensor. This sensor used to detect hand gesture and converted into code. Accuracy of that system was 90%.

c) Leap motion: Leap Motion controller is a sensor which detects the hand movement and converts that signal into computer commands. It consists of two IR cameras and three infrared LED's. LED generates IR light signal and camera generates 300 frames per second of reflected data.

These signals are sending to the computer through USB cable for further processing. The Leap device detects the data like point, wave, reach, grab which is generated by a leap motion controller. Combination of DTW and IS algorithm is used for conversion of hand gesture into text. Leap motion controller used to sense the hand movement and convert that hand movement into computer commands. Artificial neuron network is used for training symbols. The disadvantage of that system was low accuracy and fidelity.

d) Vision based: In this method web camera used to capture images. After that image segmentation has done. Feature like palm, finger extracted from input image. Different hand motion that is half closed, fully closed, semi closed was detected.

Different hand motion that is half closed, fully closed, semi closed was detected. Data is saved in vector and that vector is used for recognition of alphabets. SVM algorithm is used for classification purpose.

In this system vowels recognized with accuracy 99.4% and consonants recognized with 99.6% accuracy. Generally while capturing the image for experiments, head movement is also mixed with hand images.

To solve this overlap between hand and head movement, camera is mounted above of signers .But due to this face and body gesture lost.

1.2 Problem Statement:-

Dumb people are usually deprived of normal communication, as they find it really difficult at times to interact with normal people with their gestures, as only a very few of those are recognized by most people. Trainers usually find it difficult to teach mute community.

1.3 Architectural Design:-

The main idea is to develop and offer courses aimed at training teachers to teach ISL, develop and create resources for use in teaching of ISL to children, parents, teachers and the general public; facilitate educational use of ISL in special schools as the first language or medium of instruction and in mainstream schools as a second language or as interpreter mediated language of classroom communication.

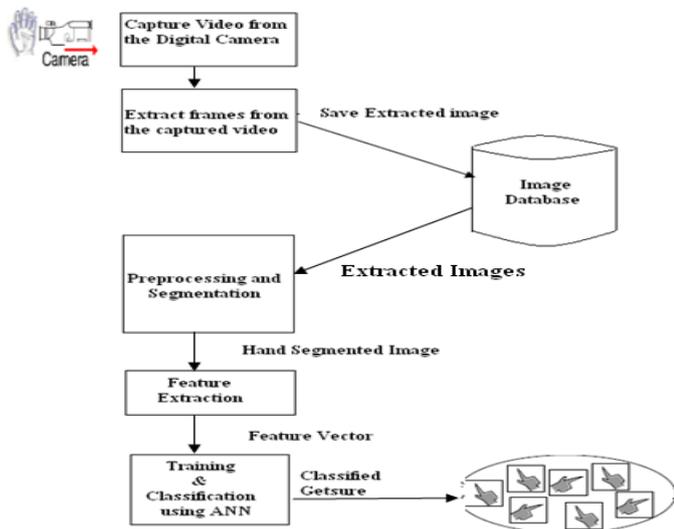


Figure : Block Diagram of Proposed System

Step 1: Capture video of hand from the digital camera. Extract the frames from the captured video. Store the video frames in image database.

Step 2: Read the frames form stored database; preprocess the image. Segment the gesture image to extract region of interest.

Step 3: Extract features from the segmented image, Store the extracted features in feature vector for training. Build the Neural network for training & recognizing of gestures.

Step 4: Test the Hand Gesture Interface.

1.4 Objectives:-

The objective of the proposed work is to build a system that uses gestures as a modality for recognition in the vision based and/or glove based setup. The focus of the proposed project is to develop a human Computer Interaction (HCI) platform in the context of Indian Sign

Language. The development of a system for translating Indian Sign Language into spoken language would be greater help for deaf as well as hearing people of our country. In a country like India there is a need of automatic sign language recognition system, which can cater the need of hearing impaired people.

2. Proposed Work:-

In order to pursuit the goals of the project the following points are identified as essential.

- a) Designing the prototype of Indian sign language education and recognition system.
- b) Gathering information regarding different regional sign languages of India.
- c) Creating skeleton of various signs and storing them into database for an interactive online environment.
- d) Capturing gestures by the use of hand gloves or camera.
- e) Making the computer understand different gestures of different sign languages and animating different gestures in real time.

System that enables dumb and deaf people to further connect with their society and aids them in overcoming communication obstacles created by the society's incapability of understanding sign language. The system we propose is based on translating gestures of the mute people into text and audio.

2.1 Proposed Modules:-

The various modules used to achieve the objectives are as follows:-

i) Camera Configure:

Capture configures mainly deals with capturing gestures of impaired people through the digital camera. Once the images are capture through camera the image is divided into frames and the frames are extracted from the captured gestures. Store the gestures frames in image database.

ii) Image Processor:

Read the extract frames from stored database and forward to pre-processor. The pre-processor process the image in which extracted capture image is converted to black white image and then to binary. The generated binary image is segmented to extract region of interest.

iii) Pattern matcher:

In pattern matcher the segmented image obtained and the stored image in database are compared and the recognized image is obtained.

iv) Convertor:

In convertor converts the recognized image obtained from pattern matcher into equivalent text and audio format.

2.2 Algorithms:-

Otsubinary Algorithm:-

A natural way to separating an image into regions is through thresholding, the separation of light and dark regions. Thresholding creates binary images from grey-

level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one.

A natural way to separating an image into regions is through thresholding, the separation of light and dark regions. Thresholding creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one.

So, for each potential threshold T we

1. Separate the pixels into two clusters according to the threshold.

If $g(x, y)$ is a threshold version of $f(x, y)$ at some global threshold T

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq T \\ 0 & \text{Otherwise} \end{cases}$$

2. Find the mean of each cluster.

We can define the within-class variance as the weighted sum of the variances of each cluster:

$$\sigma^2_{\text{Within}(T)} = n_B(T)\sigma^2_{B(T)} + n_O(T)\sigma^2_{O(T)}$$

where $n_B(T) = \sum_{i=0}^{T-1} P(i)$

$$n_O(T) = \sum_{i=T}^{N-1} P(i)$$

$\sigma^2_{B(T)}$ = the variance of the pixels in the background (below threshold) $\sigma^2_{O(T)}$ = the variance of the pixels in the foreground (above threshold) and $[0, N - 1]$ is the range of intensity levels.

3. Square the difference between the means.

Between class variance:

$$\begin{aligned} \sigma^2_{\text{Between}(T)} &= \sigma^2_{\text{Within}(T)} \\ &= n_B(T) [\mu_B(T) - \mu]^2 + n_O(T) [\mu_O(T) - \mu]^2 \end{aligned}$$

Where σ^2 is the combined variance and μ is the combined Mean.

4. Multiply by the number of pixels in one cluster times the number in the other. Substituting

$$\mu = n_B(T)\mu_B(T) + n_O(T)\mu_O(T)$$

and simplifying, we get

$$\sigma^2_{\text{Between}(T)} = n_B(T)n_O(T)[\mu_B(T) - \mu_O(T)]^2$$

Kohonen Algorithm:

The Kohonen algorithm is used for the implementation of pattern matching. Consider the problem of charting an n-dimensional space using a one dimensional chain of Kohonen units. The units are all arranged in sequence and are numbered from 1 to m. Each unit becomes the n dimensional input x and computes the corresponding excitation. The n dimensional weight vectors w_1, w_2, \dots, w_m are used for the computation. The objective of the charting process is that each unit learns to specialize on different regions of input space. When an input from such a region is fed into the network, the corresponding unit should compute the maximum excitation. Kohonen's learning algorithm is used to guarantee that this effect is achieved.

Kohonen's learning algorithm is used to guarantee that this effect is achieved

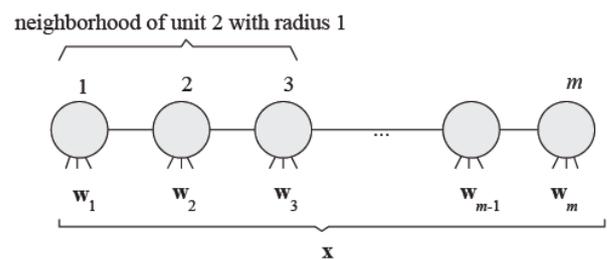


Figure:- One-Dimensional lattice of Computing Units

Algorithm:-

Start: - The n Dimensional weight vectors $w_1, w_2 \dots w_m$ of the m computing units are selected at random. An initial radius r, a Learning constant η , and a neighborhood function ϕ are selected.

step 1 : Select an input vector ξ using the desired probability distribution over the input space.

step 2 : The unit k with the maximum excitation is selected

(that is, for the distance between w_i and ξ is minimal, $i = 1, \dots, m$).

step 3 : The weight vectors are updated using the neighborhood function and the update rule.

$w_i \leftarrow w_i + \eta \varphi(i, k)(\xi - w_i)$, for $i = 1, \dots, m$.

step 4 : Stop if the maximum number of iterations has been reached; otherwise modify η and φ as scheduled and continue with step 1.

3. Scope statement and analysis:-

i) Deliverables:-

A Sign Language Recognition System for impaired people uses image processing concept to report and solve their problems.

ii) Project Deliverables:

Providing a "Sign Language Recognition System for Impaired People using Image Processing"

Sign language recognition of impaired people are usually unrecognizable by normal people. These sign languages are recognized by the system to convert it into their equivalent text and audio format.

iii) Boundary:

This system mainly deals with recognition of Indian sign language; no other languages are used for recognition.

iv) Accepting Criteria:

The images of sign languages of impaired people must be captured successfully by the camera, the captured image and the stored image must match properly and the equivalent text and audio must be generated.

v) Project Justification:

"Sign Language Recognition System for Impaired People using Image Processing" provide facilities like gesture capturing of impaired people and convert it into equivalent text and audio format. The system bridges the gap of communication between impaired and normal people.

4. Possible Outcome :-

A functioning sign language recognition system that can provide an opportunity for the deaf to communicate with non-signing people without the need for an interpreter.

5. Conclusion:-

We can use this proposed system as an interface between impaired person to normal persons who can easily understand the sign language expressed by impaired persons and also the trainers can make use of this system to train deaf and dumb persons.

References:-

- [1] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales, "Real -Time Hand Tracking and Gesture Recognition for Human-Computer Interaction" Departamento de Matemáticas e Informàtica Universitat de les Illes Balears. WSEAS Transactions on Computers, Volume 9, Issue 6 (June 2010), Pages: 583-592, Year of Publication: 2010, ISSN: 1109-2750.
- [2] Zhi Li, Ray Jarvis, "Real time Hand Gesture Recognition using a Range Camera", Australasian Conference on Robotics and Automation (ACRA), December 2-4, 2009, Sydney, Australia.
- [3] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, "Vision Based Hand Gesture Recognition", World Academy of Science, Engineering and Technology 49 2009.
- [4] G. R. S. Murthy & R. S. Jadon, "A review of Vision Based Hand Gestures Recognition", International Journal of Information Technology and Knowledge Management, July-December 2009, Volume 2, No. 2, pp. 405-410.
- [5] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang, Fellow, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", IEEE transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, July 1997.
- [6] Sushmita Mitra, Tinku Acharya, "Gesture Recognition: A Survey" IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, Vol. 37, No. 3, MAY 2007.
- [7] Hideyuki Sawada, Seiji Uta and Shuji Hashimoto, "Gesture Recognition for Human-Friendly Interface in Designer - Consumer Cooperate Design System", Proceedings of the 1999 IEEE International Workshop on Robot and Human Interaction Pisa, Italy - September 1999.
- [8] Prateem Chakraborty, Prashant Sarawgi, Ankit Mehrotra, Gaurav Agarwal, Ratika Pradhan, "Hand Gesture Recognition: A Comparative Study", Proceedings of the International MultiConferenc.
- [9] Elena Sánchez-Nielsen, "Hand Gesture Recognition for Human-Machine Interaction", Department of Statistic O.R. and Computer Science, University of La Laguna Edificio de Física y Matemáticas 38271, La Laguna, Spain.