International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 03 | Mar 2019 www.irjet.net p-ISSN: 2395-0072

Survey on Sign Language and Gesture Recognition System

Hema B N[1], Sania Anjum[2], Umme Hani[3], Vanaja P[4], Akshatha M[5]

[1,2,3,4] *Students, Dept. of Computer Science and Engineering, Vidya vikas Institute of Engineering and Technology, Mysore, Karnataka, India*

[5] *Professor, Dept. of Computer Science and Engineering, Vidya vikas Institute of Engineering and Technology, Mysore, Karnataka, India*

Abstract - *Every day we see many disabled people such as deaf, dumb and blind etc. They face difficulty to interact with others. Around 4 million of people in India are deaf and around 10 million of people in India are dumb. And they use sign language to communicate with each other and normal people. But Normal people find it difficult to understand the sign language and gesture made by deaf and dumb people. So there are many techniques which can be used to convert the sign language made by the disables into the form(such as text, voice etc) which can be understood by the normal people. Previously developed techniques are all sensors based and they didn't give the general solution. Recently number of vision based techniques are available to accomplish this. This paper explains the various available techniques that can be used to convert the sign language and gesture into the understandable format.*

***Key Words*: Sign language recognition (SLR), Human Computer Interaction (HCI), 7Hu (7 Moments), K-Nearest Neighbour (KNN), Probabilistic Neural Network (PNN), Histogram of Oriented Gradient (HOG).**

1.INTRODUCTION

Hand gesture recognition provides an intelligent, natural, and convenient way of human–computer interaction (HCI). Sign language recognition (SLR) and gesture-based control are two major applications for hand gesture recognition technologies . SLR aims to interpret sign languages automatically by a computer in order to help the deaf communicate with hearing society conveniently. Since sign language is a kind of highly structured and largely symbolic human gesture set, SLR also serves as a good basic for the development of general gesture-based HCI. [1] In this paper we are discussing work done in the area of hand gesture recognition and analyze the methods for recognition of hand gesture .Dumb people use hand signs to communicate, hence normal people face problem in recognizing their language by signs made. Hence there is a need of the systems which recognizes the different signs and conveys the information to the normal people**.** No one form of sign language is universal as it varies from region to region and country to country and a single gesture can carry a different meaning in a different part of the world. Various available sign languages are American Sign Language (ASL), British Sign Language (BSL), Turkish Sign Language (TSL), Indian Sign Language (ISL) and many more.

2. SIGN LANGUAGE

Sign language is a means of communication among the deaf and mute community. It emerges and evolves naturally within hearing impaired community. The sign language differs from people to people and from region to region. There is no particular standardized sign language they follow. As the sign language becomes a communicating language to the deaf and mute people it can be made easier for them to communicate with normal people too. Gesture recognition is an efficient way through which the movements of hand can be recognized easily. According to a survey made by some organizations there are many people who are deaf and mute and find it difficult to communicate. Gestures are made by them as an easy way of communication.

2.1 Sensor Based Approach

This approach collects the data of gesture performed by using different sensors. The data is then analyzed and conclusions are drawn in accordance with the recognition model. In case of hand gesture recognition different types of sensors were used and placed on hand, when the hand performs any gesture, the data is recorded and is then further analyzed. The first sensor used was Data gloves then LED's came into existence. The invention of the first data glove was done in 1977. Sensor based approach damages the natural motion of hand because of use of external hardware. The major disadvantage is complex gestures cannot be performed using this method. [1]

2.2 Vision Based Approach

This approach takes image from the camera as data of gesture. The vision based method mainly concentrates on captured image of gesture and extract the main feature and recognizes it. The color bands were used at the start of vision based approach. The main disadvantage of this method was the standard color should be used on the finger tips. Then use of bare hands preferred rather than the color bands. [1]

3. LITERATURE SURVEY

In this paper [2] the sign and hand gestures are captured and processed with the help of mat lab and the converted into

speech & text form .The feature extraction of values of the images is evaluated based on 7Hu(7 moments) invariant moment technique and The classification techniques are applied using K-Nearest Neighbour (KNN) is compared with the PNN(Probabilistic Neural Network) for the accuracy rate. The performance of proposed classifier KNN is decided based on various parameters Parameters can be calculate by formula's, Using 7hu moments technique's for feature extraction will be done , the 7Hu moments are a vector of algebraic invariants that added regular moment. They are invariant under change of translation, size and rotation. Hu moments have been widely used in classification [2]. And KNN (k-nearest neighbour's classification) algorithm will classifies the objects on basis of majority votes of its neighbour's. The object is assigned to the class by Most common among its nearest neighbour's k (k is a small positive integer).

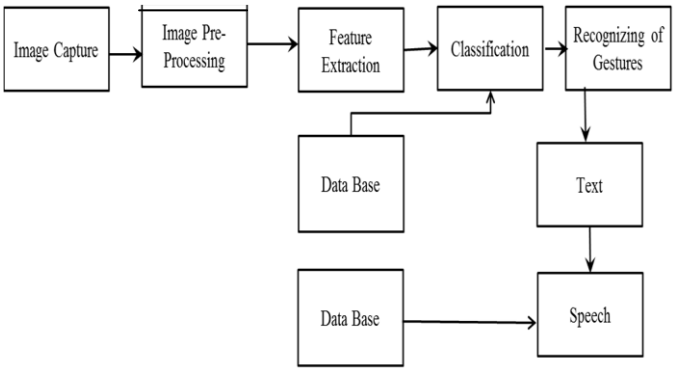

Fig-1 : Architecture of the system

Using (KNN) classification we get approximately 82% accuracy rate. The limitations of this system is the results of speech is audio output and the gestures or signs is only on the left side of the upper half of the captured image. And Capturing of image carefully and with high quality webcam is needed.

In this paper [3], hand gesture recognition can be done by wearing gloves this proposed system can work for real-time translation of Taiwanese sign language. The different hand gesture can be identified by using the 10 flex sensors and inertial sensors which is embedded into the glove, it includes three major parameters, there are

1. The posture of fingers
2. Orientation of the palm
3. Motion of the hand

As defined in Taiwanese Sign Language and it can recognized any without ambiguity. The finger flexion postures can be collected from the flex sensors, the palm orientation acquired by G-sensor, also the motion trajectory acquired by gyroscope are used as the input signals of the proposed system. The input signals will be gathered and validate or checked periodically to see if it is a valid sign language gesture or not. The flex sensor are attached with a fixed resister, to calculate the voltage between the flex sensor (*RF*) and the fixed resistor (*RB*) using below formula

$$Vo = Vi \ RF / RF + RB$$

Depending upon the values of *RF* &*RB* the angle of bent in hand figure will be obtained. The orientations of palm like up, down, right, left etc. Will be identified by using 3D data sequence along x-axis y-axis and z-axis. Once all these process done the sampled signal can last longer than a predefined clock cycles, and it is regarded as a valid hand sign language gesture and will be sent to smart phones via Bluetooth for gesture identification and to speech translation. The accuracy for gesture recognition in this proposed system is up to 94%.

In this paper [4] A vision based interface system is developed for controlling and performing various computer functions with the aim of human computer interaction. The human computer interaction aims at easy way of interaction with the system. Image segmentation and feature extraction can be done by using the vision base technology.

Human computer interaction is a multidisciplinary field of study focusing on the design of computer technology and the interaction between humans and computers. The researches done with HCI deals with the design, evaluation, implementation and assessment of new interfaces for improving the interaction between humans and machines. Here convex hull and convexity defects algorithms are programmed in Open CV. These are manipulated to various parameters which can be used for computer control. The steps involved are:

- Derive the binary image from the coloured image frames
- Find the contour of the binary image and draw this on another blank image
- Find the center of mass of hand
- Find and draw the convex hull and convexity defects on a blank image
- Define and manipulate certain points which can be used for gesture control

From the obtained colour image, binary image is derived using image segmentation technique. This is followed by extracting the contour of the image which is required for drawing the convex hull and convexity defect. The center of mass of the hand is computed using algorithms available in Open CV. The convex hull of the hand contour is drawn. The region between the fingers forms the convexity defects. The number of fingers present in the hand gesture is determined by counting the number of convexity defects. The vertices of the convex polygon are found out to identify the convexity defects. There are few parameters to be considered:

- Point start-Point of contour where the convexity defect begins
- Point end- Point of contour where the convexity defect ends
- Point far-Point within the defect which is farthest from the convex hull

- Depth-Distance between the convex hull i.e the outermost points and the farthest points within the contour.

These points obtained can be used as parameters for designing hand gestures for computer control.

In this paper [5], a vision based sign language recognition system using LABVIEW for automatic sign language has been presented. The goal of this project is to develop a new vision based technology for recognizing and translating continuous sign language to text. Although the deaf, hard of hearing and hearing signers can communicate without problems among themselves ,there is a serious challenge for deaf community trying to integrate into educational ,social and work environments.

Methodology:

Vision based sign language translation is a real-time system which is broken down into three main parts starting with Image acquisition followed by image processing and last comes image recognition. The program starts with image acquisition .i.e., sign images capturing by the camera. The acquired images are pre-processed to differentiate static and dynamic signs, and also the start and end of the sign. The images are extracted to be used in the recognition stage. In the recognition stage, the features extracted are compared with the available database of pattern matching templates.

A threshold value is set for the maximum difference between the input sign and the database, if the difference is below the maximum limit, a match is found and the sign is recognized.

Environmental setup:

Image acquisition process is subjected to many environmental concerns such as the position of the camera, lighting sensitivity and background condition.

Result:

This sign language translator is able to translate alphabets (A-Z/a-z) and numbers (0-9). All the signs can be translated real-time. The current system has only been trained on a very small database. This is a wearable phone system that provides the greatest utility for automatic sign language to spoken English translator. It can be worn by the signer whenever communication with a non-signer might be necessary

In this paper [6] there are two feature descriptors used Histograms of Oriented Gradients and Scale Invariant Feature Transform. The algorithm implemented in KNN uses HOG features for an image and classifies them using SVM technique. In the proposed system the algorithm implemented to recognize the gestures of Indian Sign Language is of static images.

- Categorization of single handed or double handed: sub categorization is the main aim here which is implemented using HOG-SVM classifier. Here the images are converted from colour to binary. Hence SVM is quite accurate in separating single and double handed
- Feature extraction: for both single and double handed feature recognition HOG and SIFT are extracted and these two descriptors are used. The HOG uses histogram to depict the intensity distribution which is done by dividing the image into cells and creating 1D histogram for edge orientation in each cell. In SIFT, initially the points of interest are called as key points, are located which represents local maxima and minima that are applied to the whole image. Next the key points are filtered by eliminating low contrast points by setting a threshold
- Feature fusion using K-nearest correlated neighbour features algorithm: the matrices of HOG and SIFT features for the training images are constructed separately for single handed and double handed features. The correlation of SIFT an SIFT for a test image is calculated with all other features of the dataset

Result:

Without categorising into single and double handed gestures:

- Accuracy using KNN algorithm of HOG features of images is 78.84% and that for SIFT gave 80%.
- After classifying into single and double handed gestures:
- HOG gave100% for single handed 82.77% for double handed
- SIFT gave 92.50% for single handed 75.55% for double handed
- Fusion of both gave 97.50% for single handed and 91.11% for double handed
- Validation of accuracy of algorithm: the proposed method was able to identify 59 test images correctly out of 60 test images for single handed gestures. For double handed gestures out of 200 test images it was correctly able to identify 179 of them

In the today's world there are many disabled people (deaf, mute, blind, etc) who face lot of problems when they try to communicate with other. Previously developed devices did not implement any general solution. This paper [7] describes a new method of developing wearable sensor gloves for detecting hand gestures which uses British and Indian sign language system. The outputs are produced in the text format using LCD and audio format using APR9600 module. The hand gesture or the hand signs are converted to electrical signals using flex sensor. These electrical signals are processed to produce appropriate audio and text output.

Previously designed devices were not accurate in tracing the hand gestures. The paper employs method of tuning in order to improve the accuracy of detecting hand gesture. The paper as gives the statistical overview of the population of India in both urban and ruler area suffering from at least one disability . It discuss in brief the type of gesture recognition available the gesture recognition can be done in three ways, i.e., vision based, sensor based and hybrid of the fit two They have explained how the processing has been divided into three stage that is sensors ,processing unit and output

A glove with flex sensor is used to detect the signs made by the person. The made. Gestures which are made are converted to equivalent electrical signals by the sensor. These signals are sent to input pins of the MSP430F149. For different gestures the signals are different. A range of ADC values [12bit ADC assumed] is assigned for each gesture the corresponding word would be stored in the memory. When the controller receives these signals it compares with the stored values and the corresponding word is selected and the word is spoken out of the speaker. Speaker here can used. The corresponding word is simultaneously displayed on the LCD. The complete block diagram is shown in Figure below,

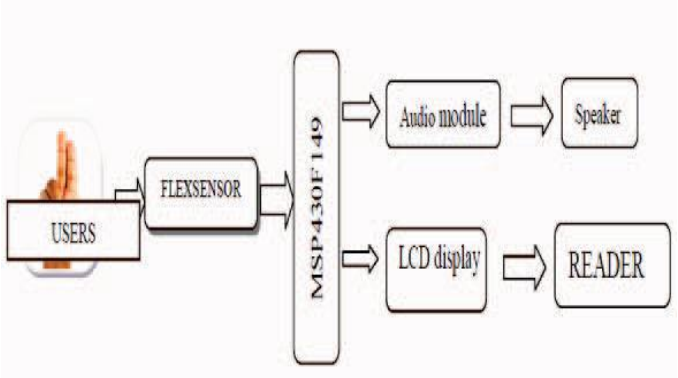

Fig-2: Block diagram of working model

This paper [7] also explain about the various hardware devices like **Gloves, MSP430F149, APR9600** and their functionality in the proposed system. The working of software how the gesture is converted into electrical signal and forwarded to the processing unit. The advantage of using gloves for in verbal communication is the are cheap and accurate. The drawback is processing the delay.

This paper [8] tells about the first data driven automatic sign-language to speech translation system. They have combined a SL (sign-language) recognizing framework with a state-of-art phrased based machine translation (MT) system using corpora (complete written work) of both American sign language and Irish sign language data. They have also explained about vision-based knowledge source. They also inform us about the challenges that comes down when a deaf and hearing person communicate. They have suggested that data-driven output recognizer is not easily intelligible because of different grammar and annotation format

Sign Language: Sign language is local to specific countries SL are grammatically distinct from the spoken language and grammar and it make use of visual/gestural modality hand are used as information carrier it also make use of facial expression tilts of the heads and shoulders as well as velocity of sign

Sign Language Transcription: There is lack of formally adopted writing system of SL some attempts have been made to describe the hand shape location and articulation moment despite this they fall short to be used in computational hence they adopt annotation ,in annotation manual transcription of sign language is taken from video data.

The Corpus: Data-driven approach MT require a bilingual data set (speaking two language fluently)

Sign language recognition: The automatic sign language recognition (ASLR) system is based on The automatic speech recognition (ASR) system. To model image variability (no fixed pattern) various approaches are known and have been applied to gesture recognition**.** To extract features the dominant hand is tracked in each image sequence hence robust tracking algorithm is required Data-driven SL MT**.** This area is new they are 4 method employed in this approach

1. Example –based MT for English sign language of the Netherland

2. SMT (statistical machine translation) for German and German sign language

3. (chiu) a system for Chinese and Taiwanese sign language

4. (san-segundo) for Spanish and Spanish sign language

SMT : SMT uses a state-of –the-art phrase based SMT system to automatically transfer the meaning of source language sentence into a target language sentence

Result:

The size of RWTH-Boston 104 is far too small to make reliable assumption at very least we show that SMT is capable to work as an intermediate step for complete sign-to-speech system. For extremely small training data the resulting translation quality is reasonable. As no sign language parser exits for annotation they propose stemming of the glosses (leaving out the flexion) during the recognition. Moreover sign language produce quite a large effect known as co articulation (the movement between two sign that cannot be easily trained) .

This paper [9] is about to design a system that can understand the sign language and gestures accurately so that the deaf and dumb people may communicate with the outside world easily without the need of an interpreter to help them .

Here the work is done for the American sign language (ASL) which has its own unique gesture to represents the words and the English alphabets i.e. is from A-Z. The process is to recognize the sign language and gesture as shown in the below figure

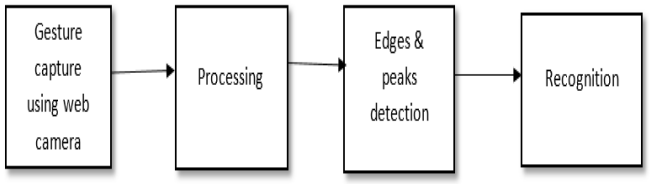

Fig-3: System block diagram

The image will be captured through the web camera and it stored in system. The acquired image that is stored in the system windows needs to be connected to the mat lab software automatically to continue the further process. This is can be done by creating an object s of that images. With the help of high speed processors in system .The imaqe info function will be used here to determine if the device supports device configuration files or not. If the input is an RGB image, it can be of class uint8, uint16, single, or double. The output image is of the same class as the input image. These captured colour (RGB) image first converted into grayscale or binary image as some of the pre-processing techniques can be applied on grayscale and binary image only. Further these binary images under goes the canny edge detection technique .The Edge Detection block finds the edges in an input image by approximating the gradient magnitude of the image. The block convolves the input matrix with the Sobel, Prewitt, or Roberts kernel. The block outputs two gradient components of the image, which are the result of this convolution operation and the output a binary image, is a matrix of Boolean values that is 0 and 1's forms. If a pixel value is 1, it is an edge or else it's not edge part of image. A KNN (k nearest neighbors) algorithm used the classification purpose and alternatively Watershed transform process will be used instead of KNN which computes a label matrix which can identify the watershed regions of the input matrix, which can have any dimension. And the elements of L are integer values greater than or equal to 0. The elements labelled 0 do not belong to a unique watershed region, and at the end figure tip is detected. Finally the last step is recognition which match the gesture with the predefined gesture database and produces the suitable matched meaning of the gesture in speech and text form.

4. CONCLUSIONS

This paper deals with the different algorithms and the techniques that can be used for recognizing the sign language and the hand gesture made by different deaf and dumb people. Hand gesture recognition system is considered as a way for more intuitive and proficient human computer interaction tool. The range of applications includes virtual prototyping, sign language analysis and medical training. Sign language is one of the tool of communication for physically impaired, deaf and dumb people. From the above consideration it is clear that the vision based hand gesture recognition has made remarkable progress in the field of hand gesture recognition. Software tools which are used for implementing the gesture recognition system are C, C++, and Java language. To simplify the work especially when image processing operations are needed, MATLAB with image processing toolbox is used.

REFERENCES

[1] IJTRS-V2-I7-005 Volume 2 Issue VII, August 2017 survey on hand gesture techniques for sign language recognition

[2] Umang Patel & Aarti G. Ambekar "Moment based sign language recognition for Indian Languages "2017 Third International Conference on Computing, Communication, Control and Automation (ICCUBEA)

[3] Lih-Jen kau, Wan-Lin Su, Pei-Ju Yu, Sin-Jhan Wei "A Real-time Portable Sign Language Translation System" Department of Electronic Engineering, National Taipei University of Technology No.1, Sec. 3, Chung-Hsiao E. Rd., Taipei 10608, Taiwan, R.O.C.

[4] Obtaining hand gesture parameters using Image Processing by Alisha Pradhan and B.B.V.L. Deepak ,2015 International Conference on Smart Technology and Management (ICSTM).

[5] Vision based sign language translation device (conference paper: February 2013 by Yellapu Madhuri and Anburajan Mariamichael).

[6] K-nearest correlated neighbour classification for Indian sign language gesture recognition using feature extraction(by Bhumika Gupta, Pushkar Shukla and Ankush Mittal).

[7] Vikram Sharma M, Virtual Talk for deaf, mute, blind and normal humans, Texas instruments India Educator's conference,2013.

[8] Hand in Hand: Automatic Sign Language to English Translation, by Daniel Stein, Philippe Dreuw, Hermann Ney and Sara Morrissey, Andy Way

[9] Er. Aditi Kalsh, Dr. N.S. Garewal "Sign Language Recognition System" International Journal of Computational Engineering Research 2013

[10] https://en.wikipedia.org/wiki/Sign_language