© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

A Survey of Sign Recognition Approaches for Indian Sign Language

Twinkal H. Panchal[1], Pradip R. Patel[2]

[1]M.E. Student (Information Technology Department), L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

[2]Assistant Professor (Information Technology Department), L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Sign Language (SL) is a communication tool for deaf & dumb people. It is a subset of gestures made with fingers, hands, and face etc. Each gesture in SL has a particular meaning & that is assigned to it. Deaf person directly not communicate with the normal person because normal person never try to learn the sign language. To solve this problem, there exists a need of system that can recognize gesture. Different country has different sign languages. For India, this is called as "Indian Sign Language (ISL)". Only little research work has been carried out in this area for ISL. Several methods have been used to recognize of ISL alphabets and numerals. Many of them method is used to recognize static gesture. Only few works have been reported for dynamic gesture recognition of ISL. The mainly four steps are involved to recognize the sign: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. This paper presents a survey on various sign recognition approaches for ISL.

Keywords: Indian Sign Language, Gesture acquisition, Hand Gesture Recognition, Segmentation, Feature Extraction

I. INTRODUCTION

Sign language is the only and cheapest approach of communication for the people who is suffering from disability like hearing and speaking. According to the survey taken by government of India, it is reported that in year 2011 census data over 2.68 crore of people in India suffer from some form of the disability. Out of this 18.9% people have speech and 7.5% people have hearing disability [1]. Sign language is composed of visual gestures and signs where every sign has a specific meaning allotted to it. There are 143 existing different sign languages all over the world, mainly American Sign Language, British Sign Language, French Sign Language, Japanese Sign Language, and Indian Sign Language [2]. The gestures are mainly divided into two types: Static gestures and Dynamic Gestures. Static gestures include only poses and configurations whereas dynamic gestures include strokes, pre-strokes, postures and phases. The dynamic gestures include movement of body parts. It may also include emotions depending on the meaning that gesture conveys [6].

ISL alphabets and numeric signs [8] are represented in Figure 1 and Figure 2 respectively.

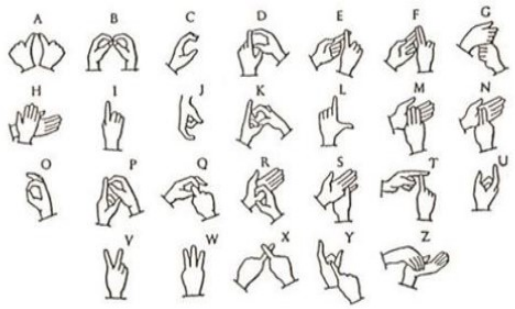

Figure 1. Representation of ISL Alphabets.

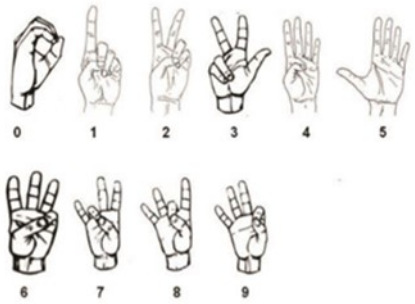

Figure 2. Representation of ISL numerals.

To make communication easy between deaf-dumb and normal people, sign recognition system that first recognize the gestures and convert it into text or voice. Developing such sign recognition system for ISL is more challenging than other sign languages due to the following reasons:

- ✓ In ISL both static & dynamic gestures are used.
- ✓ One hand move faster than other in Dynamic hand Gesture.
- ✓ ISL include facial expression.
- ✓ Complex hand shapes.

A very few work has been carried out to develop a system which converts ISL into text and voice because of these challenges. In this paper, various method used for hand gesture recognition system in ISL.

II. LITERATURE SURVEY

Surbhi Rathi and Ujwalla Gawande proposed a full duplex communication system for deaf & dumb people [5].The proposed system for recognizing a dynamic hand words gesture of Indian sign language and conversion of recognized gesture into text and speech and vice versa i.e. dual way communication. Pre-processing is divided into two steps: Segmentation & Filtering. In this system skin color filtering technique has used for segmentation. Eigen vectors and Eigen values technique has used for feature extraction. For classification, Eigen value weighted Euclidean Distance based classifier has used. The proposed system used minimum Eigen Value distance for recognition of sign. In dual way communication system first the recognized gesture has converted into text message and voice format, so that normal person can understand it. Second, voice has converted into text and its corresponding gesture, so that physically impaired humans can understand it.

Purva C. Badhe and Vaishali Kulkarni proposed a system for dynamic recognition of ISL gestures [6]. ISL alphabets & numerals convert into text in the proposed system. For tracking hand movement they used a combinational algorithm which included canny edge detection, skin color detection with YCbCr, thresholding, etc. For feature extraction they used Fourier descriptors method. In this system template matching is used to recognize sign. Self-created database is used for implementation. The proposed system gives an accuracy of 97.5%. MATLAB is used to creating a Graphical User Interface (GUI) for developing this system.

Bhumika Gupta, Pushkar Shukla, Ankush Mittal proposed a system for recognize a static images of the signed alphabets in the Indian Sign Language[7].This system first categorize sign into single handed or double handed. They used Histograms of Oriented gradients(HOG) Descriptor and Scale Invariant Feature Transform(SIFT) Descriptor for feature extraction. They used feature Fusion using K-Nearest Correlated Neighbour Algorithm for classification. The dataset constitutes of 520 images for the training segment and 260 images for testing. Fusion of both kinds of descriptors gave an accuracy of 90% without categorization into single handed &double handed gesture. Fusion of both kinds of descriptors gave an accuracy of 97.50% for single handed gestures and 91.11% for double handed gestures.

Aditya V., Vinod P.R., Usha Gopalkrishnan proposed a system based on Artificial Neural Network[8]. They proposed system recognition the static gestures in ISL alphabets and numerals.ISL fingerspelling translate into text in this system. The gestures included English alphabets (26 letters) and numerals (0-9). YCbCr model, filtering and morphological operations for hand segmentation is used in this system. For classification they used artificial neural network. 36 signs with 15 images of each were tested in this system. The system was implemented using MATLABR2010a.Low Computational capacity in this system. This proposed System gives an accuracy of 91.11%.

Shreyashi Narayan Sawant, M.S.Kumbhar proposed a system for real time sign recognition using PCA[9]. For Segmentation they used Otsu algorithm. 10 each of the 26signs total 260 images consisting in dataset.At a resolution of 380 x 420 pixels the image were captured. In this method test and train image Euclidean distance was calculated and gestures have a minimum Euclidean distance. Gesture was transformed into text and voice format.

The proposed system in [10] is used for dynamic hand gesture recognition in which a key frame extraction technique is used. This algorithm is used for to find out the most important frames from the video, based on the change in hand shape and position using certain

parameters and dynamic threshold. For Pre-processing they used YCbCr model. The main parameter to be considered are hand shape, hand motion and hand orientation for feature extraction. Multiclass Support Vector Machine (MSVM) is used for classification. The use of key frame extraction algorithm speeds up the system by selecting essential frames and eliminating extra frames**.** The proposed system gives accuracy of 90.46%.

Joyeeta Singha, Karen Das proposed a system using Eigen value weighted Euclidean distances as a classification technique for recognition sign alphabets of ISL[4].They Proposed system for static sign alphabets of ISL.For Pre-processing they used Skin filtering & hand cropping. Eigen value &Eigen vector is used for feature extraction. They used two level of classification. First, Classification based on Euclidean Distance. Second, Classification based on Eigen value weighted Euclidean distances. The proposed system gives accuracy of 97%.

III. DIFFERENT TECHNIQUES FOR RECOGNITION

Gesture recognition can be divided into two types: sensor based and vision based. In sensor based method, the gestural data can be extracted using data glove or motion sensors [6]. Data glove captured Even minute details of the gesture which ultimately enhances the system performance. Wearing a hand glove with embedded sensors in this method [6]. This method also reduces the user comfort and it affects the signer's usual signing ability.

Image processing include in vision based method. This approach is comfortable to the user because no extra device wearing. With the help of camera images is captured. For identifying the gesture this method deals with the features of image such as color and texture. Complexity of the background, variation in illumination, and tracking of other skin color objects along with the hand objects are the challenges in vision based approach.

Figure 3. (A) Data-Glove based[11]

Figure 3. (B) Vision based[11]

IV. OVERVIEW OF SIGN LANGUAGE RECOGNITION SYSTEM

The mainly four steps are involved in a sign language recognition system are: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. First step of sign recognition system is to acquire gestural data. A webcam is used to capture the images. Segmentation is required to tracking the hand movement. To extract important features feature extraction is used. The process of recognition can be divided into the two stages - training and testing. In training stage the classifier is trained using the training database. The main steps in training stage are creation of database, pre-processing, extraction of features and training of the classifier. The main steps involved in the testing phase are gesture acquisition, preprocessing, feature extraction and classification. Fig.4 shows block diagram of sign language recognition system.

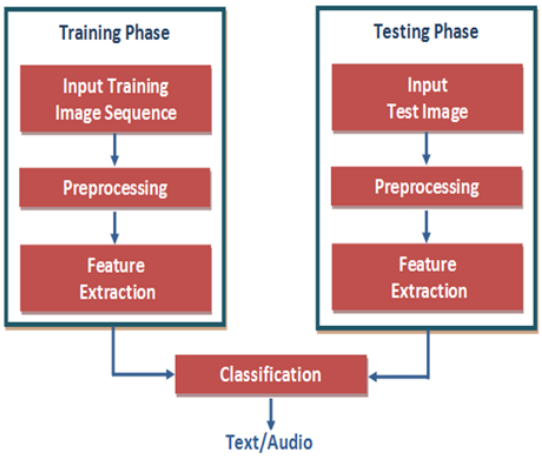

Figure 4. Sign Language Recognition System

A. Pre-processing

Preprocessing step is performed before feature extraction to produce the most useful information neglecting meaningless information. It consists of image enhancement, image filtering, image resizing, and image segmentation. To extract the region of interest the

images or videos of the test database are also preprocessed during testing phase.

In [5] skin color filtering technique is used for preprocessing. First the video is converted into frames.Then, pre-processing techniques has applied to the frames which are divided in to two parts namely filtering and segmentation. Mostly, skin color filtering technique is applied on input video frames for the detection of hand gesture from the background of the image. Hence, skin color filtering technique separates skin colored region from non-skin colored region. The entire frames are converted to HSV color space. Hue, Saturation and Value color model is used for skin detection and also it is less sensitive to illumination changes in comparison of RGB.

Hence, it has shown to be one of the most adapted techniques for skin color detection. Then, to get the rid of high frequency noise from the input image Gaussian low pass filter has applied, which has also used for the smoothning of image frames. Segmentation is used to find the only hand region from the entire image region. After removing noise from the HSV image convert this image in to gray scale image first and then binary image. So, the result obtained is a binary image with skin region is in white and non-skin region is in black color.

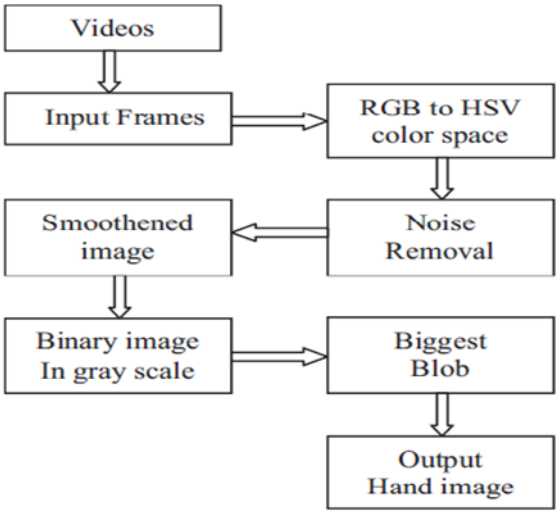

Figure 5. Block Diagram of skin color filtering

In[6] Skin color detection with YCbCr and Canny Edge detection are used for preprocessing. In this, video is first converted into frame. After cropping the frames, two subsequent frames are subtracted from each other for obtaining the difference image. Calculate the threshold & OR with the difference image. They used YCbCr for Skin color detection & Canny Edge Detection for boundary Detection. Then all are combined in one frame. Then after send it to feature extraction step.

In [7] , images are converted from a colored to a binary format by the employment of RGB thresholding . For removal of noise, the images are extracted for their HOG features. These features when fed to a standard Support Vector Machine were quite accurately able to separate single-handed gestures from the double-handed.

In [8], YCbCr color space is used for preprocessing & Hand segmentation. In order to detect the skin color in the input image it is first converted to YCbCr color space. YCbCr separates RGB into luminance and chrominance components where Y is the luminance component and Cb, Cr are the chrominance components. RGB values can be transformed to YCbCr color space using the following equations.

$$Y = 0.299R + 0.587G + 0.114B,$$
$$Cr = 128 + 0.5R - 0.418G - 0.081B,$$
$$Cb = 128 - 0.168R - 0.331G + 0.5B.$$

Thus a pixel is classified as belonging to skin if it satisfies the following relation:
$75 < Cb < 135$ and $130 < Cr < 180$ and $Y > 80$.

The result of segmentation produces a binary image with the skin pixels in white color and background in black color. The resulting binary image may contain noise and segmentation errors. Filtering and morphological operations are performed on the input image to decrease noise and segmentation errors.

In [9], the Segmentation of hands is carried out to separate object. For segmentation Otsu algorithm is used. The segmented hand image is represented certain features. These features are further used for gesture recognition To remove noises from images morphological filtering techniques are used. The preprocessing operation is done on the stored database.

B. Feature Extraction

Feature extraction stage is used to extract certain features from the hand image which is unique for each sign. Feature extraction is a method of reducing data dimensionality by removing less discriminative data. The selection of appropriate features and feature extraction methods is the most significant decision in the

design of gesture recognition system.In this step feature vector is obtained.In the next step feature vector is given. Various feature extraction techniques are Fourier descriptors, HOG features, PCA, etc.

In [5], Feature extraction stage is used to taken out certain features from the hand image which is unique for each sign . If the gesture gets match, then the last frame have chosen as a reference frame and the features like Eigen vectors and Eigen values have extracted from the reference frame. Following are the mathematical steps used for calculating Eigen values and Eigen vectors:

Step 1: Let, the last frame is assumed as 'X', 'X' has to be change the size using dimension m*m.

Step 2: Calculate the mean 'M' of the above vector X as,

$$\mathbf{M = E\{X\}} \quad (2)$$

Step 3: Calculate the covariance 'C' as given in eq. (1),

$$\mathbf{C = E\ \{(X - M)\ (X - M)\ '\}} \quad (3)$$

Step 4: From the above covariance 'C' calculate the Eigen values and Eigen vectors and arrange the Eigen vectors in descending order according to the Eigen values.

Step 5: Take only first 5 principles vectors from total m Eigen vectors that reduces the dimension of the matrix without much loss of information.

In [8],The distance transform of the image is used for feature extraction and recognition tasks. In the proposed method, the distance transform is computed by using the Euclidean distance. The equations for computing the distance between two pixels with coordinates (x, y) and (u, v) are shown below:

The city-block distance between two points P = (x, y) and Q = (u, v) is defined as

$$\mathbf{d\ (P,Q) = |\ x - u\ |\ +\ |\ y - v\ |.}$$

The chessboard distance between P and Q is defined as

$$\mathbf{d\ (\ P,\ Q\ ) = max(|\ x - u\ |\ ,\ |\ y - v\ |).}$$

The Euclidean distance between P and Q is defined as

$$\mathbf{d\ (\ P,\ Q\ ) = } \sqrt{(x-u)^2 + (y-v)^2}\mathbf{.}$$

In [7] Histogram of Oriented Gradients and Scale Invariant Feature Transform descriptors are used for feature extraction.

Histograms of Oriented Gradients (HOG) Descriptor:

A mainly used descriptor for detection of an object in an image and even its shape or color which was first given by Dalal and Triggs for Pedestrian detection in. First dividing the image into smaller units called cells and constructing single dimensional histograms for the edge orientations of pixels in each cell. In order to attain invariance from varying magnitudes of illumination, these local histograms are normalized for a group of cells which is referred to as a block. These blocks of normalized histograms are used as descriptors for an image.

Scale Invariant Feature Transform (SIFT) Descriptor:

This set of descriptors defined for an image is invariant to varying illumination as well as rotation and scaling. To define the SIFT features of an image generally constitutes of four stages. SIFT can be used both as a feature detector and a descriptor. First, the points of interest, referred to as key points, in an image are located for a scale-space. These are the locations in this scale-space that represent the local maxima or minima obtained when a set of Difference of Gaussian (DoG) filters are applied all over the image.

Table 1. Feature Extraction Techniques

Methods	Literature availability
Principle Component Analysis [9]	Occluded and overlapped gestures can be recognized ,Not scale invariant
Fourier Descriptors [6]	Simple, robust, computationally efficient and immune to rotation or Scaling of shape and noise.
HOG descriptors [7]	Invariant to illumination change and orientation ,Not scale invariant

C. Classification

The feature vector is used as the input of the classifier that recognizes the sign. Classification involves two phases: training phase and testing phase. The feature vectors are given to the classifier in training phase. Identification of classes is taken in testing phase, test image or video is given as an input and the result is produced as a text and then in voice. Classifier used are Euclidean distance [5], K Nearest Neighbor (KNN) [7], Artificial Neural Network [8], Support Vector Machines (SVM) [7]etc. The classifier's performance is measured in terms of rate of recognition.

Table 2. Recognition Techniques

Methods	Literature Availability	Acc%
Support Vector Machine [7]	Accurate results robust to noise, Computationally expensive	94.23
Euclidean distance [5]	Invariant to rotation ,Time Consuming	97.50
Artificial Neural Network [8]	Low computational complexity, Time Consuming	91.11
K Nearest Neighbor (KNN) [7]	Robust to noisy data ,to determine value of parameter K	80.00

V. CONCLUSION

This paper presents a survey on various sign recognition approaches for Indian Sign Language. The main aim of the sign recognition is remove the communication gap between deaf-dumb & normal people. More work have been carried out in static sign recognition. Only few work has been carried out for dynamic hand gesture .This paper present comparison between various recognition technique. In future there is possibility to do more work in dynamic sign, word ,sentences of ISL.

VI.REFERENCES

[1] Website: Http: // censusindia.gov.in / Census_And_You/ disabled_population.aspx.

[2] Dutta, KusumikaKrori, "Double handed Indian Sign Language to speech and text." 2015 Third International Conference on Image Information Processing (ICIIP). IEEE, 2015. pp. 374-377.

[3] AditiKalsh, N.S. Garewal," Sign Language Recognition System for Deaf and Dumb", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue. 9, pp. 103-106.

[4] JoyeetaSingha , Karen das , "Indian Sign Language Recognition using Eigen Value Weighted Euclidean Distance based Classification Technique",2013 International Journal of Advanced Computer Science and Applications(IJACSA),2013.

[5] SurbhiRathi and UjwallaGawande, "Development of Full Duplex Intelligent Communication System for Deaf and Dumb People," 2017 IEEE7th International Conference on Cloud Computing, Data Science & Engineering – Confluence ,pp.733-738.

[6] P. C. Badhe and V. Kulkarni, "Indian sign language translator using gesture recognition algorithm," 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), Bhubaneswar, 2015, pp. 195-200.

[7] Bhumika Gupta, PushkarShukla ,Ankush Mittal , "K-Nearest Correlated Neighbour Classification for Indian Sign Language Gesture Recognition using Feature Fusion", 2016 International Conference on Computer Communication and Informatics(ICCCI),2016.

[8] V. Adithya, P. R. Vinod and U. Gopalakrishnan, "Artificial neural network based method for Indian sign language recognition," Information & Communication Technologies (ICT), 2013 IEEE Conference on, JeJu Island, 2013, pp. 1080-1085.

[9] Shreyashi Narayan Sawant and M. S. Kumbhar ,"Real Time Sign Language Recognition using PCA", 2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (lCACCCT),2014.

[10] BhumikaPathak, Subhash Chand Agrawal , Anand Singh Jalal, CharulBhatnagar , " A Framework For Dynamic Hand Gesture Recognition Using Key Frames Extraction" 2015 IEEE.

[11] Juhi Ekbote,Mahasweta joshi, "A survey on Hand gesture recognition for Indian Sign Language"International Research journal of Engineering and Technology (IRJET) 2016.