International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 10 | Oct -2016 www.irjet.net p-ISSN: 2395-0072

A Survey on Hand Gesture Recognition for Indian Sign Language

Miss. Juhi Ekbote[1], Mrs. Mahasweta Joshi[2]

[1]*Final Year Student of M.E. (Computer Engineering), B.V.M Engineering College, Vallabh Vidyanagar, Gujarat, India,*

[2]*Assistant Professor, B.V.M Engineering College, Vallabh Vidyanagar, Gujarat, India*

Abstract - *Sign Language is most accepted and meaningful way of communication for deaf and dumb people of the society. Sign language uses gestures, head/body movements and facial expressions for communication. It is a powerful means of Non-Verbal communication among humans. Every country has its own developed Sign Language. The language which is used in India is called as "Indian Sign Language (ISL)". Only little research work has been carried out in this area as ISL has got standardized recently. Currently many researchers have mainly focused on gesture recognition that has been recorded under static hand gesture. Only few works have been reported for recognition of dynamic hand gesture. Many methods have been developed to recognize alphabets and numerals of ISL. The major steps involved in designing such a system are: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. This paper presents a survey on various hand gesture recognition approaches for ISL.*

Key Words*: *Indian Sign Language, Hand Gesture Recognition, Segmentation, Feature Extraction, Non-Verbal Communication, Hearing and Speech Impaired

1. INTRODUCTION

Sign language provides a way for the deaf and dumb people to communicate with the outside world. Instead of voice, sign language uses gestures to communicate. Sign language is an organized way of communication in which every word or alphabet is assigned to a particular gesture [2]. Sign language is made up of a range of gestures produced by various facial expressions and movements of hands or head/body. In the last several years there has been an increased interest among the researchers in the field of sign language recognition to introduce a means of interaction from human –human to human – computer interaction. Various applications of Gesture Recognition System are: Human Computer Interface, Video gaming, Augmented reality, Home appliances, Robotics, Sign Language, etc. Hearing-impaired and speech impaired people depend on sign language interpreters for communication. But finding skilled and trained interpreters for their day to day interactions during life time is a very difficult job and also too expensive [7]. All over the world, a variety of sign languages exist. The sign language depends on the traditions and spoken language of that place. Indian sign language (ISL) is used by hearing and speech impaired people in India [3]. The gestures are primarily divided into two classes: Static gestures and Dynamic Gestures. Static gestures include only configurations and poses whereas dynamic gestures include postures, strokes, prestrokes, phases and also emotions [1]. ISL alphabets and numeric signs are represented in Fig.1 and Fig.2 respectively.

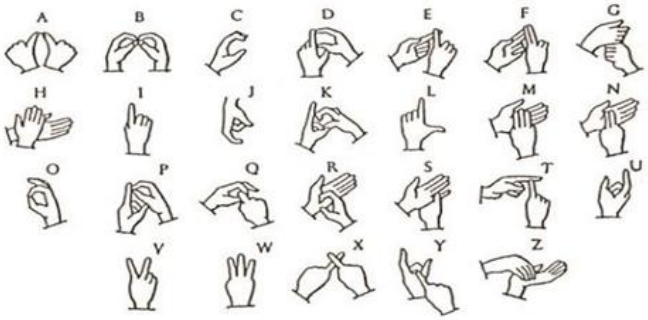

Fig.-1: Representation of ISL Alphabets (A-Z) [3]

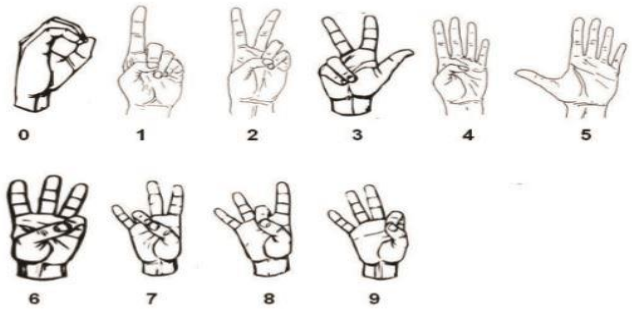

Fig.-2: Representation of ISL numerals (0-9) [3]

ISL is a standard language which provides a powerful means of communication between normal people and deaf and dumb people. Using hand gesture recognition system, one can decode the gestures and identify the signs and final output is given as a text or voice. Developing an automatic hand gesture recognition system for ISL is more challenging than other different sign languages due to the following reasons:

- Both single hand and double hand gestures are used to represent most of the alphabets.
- Both static and dynamic hand gestures are used in ISL.
- In dynamic hand gestures, one hand moves faster than the other.
- Facial expressions are also included in ISL.
- Complex hand shapes.
- Head/Body postures.

Because of these challenges, a very minute research work has been carried out to develop a system which converts ISL into resultant text and voice. Considering all the points, this survey presents different techniques for hand gesture recognition system in ISL.

2. DIFFERENT TECHNIQUES FOR RECOGNITION OF SIGN LANGUAGE

A vital application of gesture recognition is sign language recognition. Current technologies for gesture recognition can be divided into two: sensor based and vision based. In sensor based method, data glove or motion sensors are included from which the gestural data can be extracted [1]. Even minute details of the gesture can be captured by the data glove which ultimately enhances the system performance. However this method requires wearing a hand glove with embedded sensors which makes it a cumbersome device to carry [1]. This method affects the signer's usual signing ability and it also reduces user comfort. Vision based method includes image processing. This approach provides a comfortable experience to the user. The image is captured with the help of camera(s). No extra devices are needed in vision based approach. This method deals with the features of image such as color and texture that are mandatory for identifying the gesture. Although, vision based approach is straightforward but it has many challenges such as the complexity of the background, variation in illumination, and tracking of other skin color objects along with the hand object, etc. are raised. Fig. 3 shows an example.

Fig.-3: (A) Data-Glove based

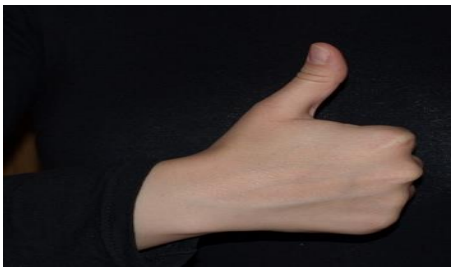

Fig.-3: (B) Vision based

3. OVERVIEW OF SIGN LANGUAGE RECOGNITION SYSTEM

The major steps involved in designing a sign language recognition system are: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. Primary step of gesture recognition system is to acquire gestural data. A webcam integrated in the laptop or an external webcam or a simple digital camera can be used to capture the images. Either an on hand database can be used or it can be formed by the researchers themselves. To track the movement of hand, segmentation of hand is required. Segmentation separates ROI from background. Hand tracking is a technique used to know position of hand. Next step is feature extraction which is carried out to extract important features after successful completion of hand tracking and segmentation. The complete process of recognition can be separated into the following two stages - training and testing. Training is the initial stage where the classifier is trained using the training database. The key steps implicated in training stage are creation of database, pre-processing, extraction of features and training of the classifier. The key steps involved in the testing phase are gesture acquisition, preprocessing, feature extraction and classification. Fig.4 shows block diagram [2][5] of sign language recognition system.

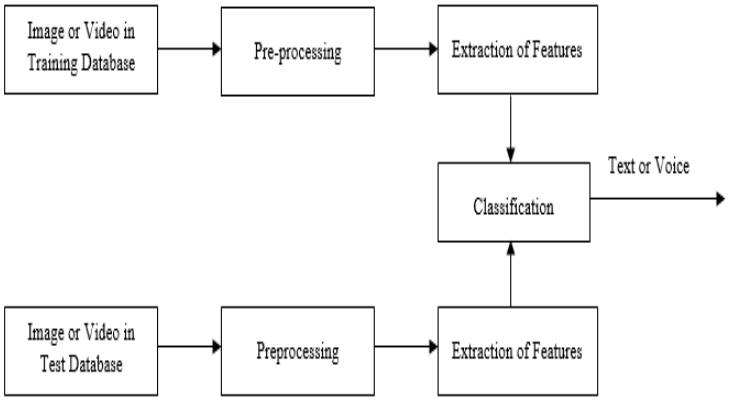

Fig.-4: Basic Steps of Sign Language Recognition System

3.1 Pre-processing

A pre-processing step is the first step to be carried out on the images which are stored in training dataset in order to extract the region of interest (ROI). This step is performed before extracting features in order to produce the most useful information neglecting meaningless noise, redundant and superficial information. The region of interest can be hand object if merely hand gestures are taken into account. Usually the pre-processing step consists of enhancement of image, filtering, resizing of image, image segmentation and morphological filtering. Commonly used segmentation techniques are Sobel edge detection [2], Otsu's thresholding [5], skin color based segmentation [3], motion based segmentation, background subtraction, Camshift method [9], HSV color model [7] etc. The images or videos of the test database are also preprocessed during testing phase to extract the region of interest.

TABLE-1: Segmentation and Tracking Techniques

Methods	Literature Availability	Remarks
HSV color model [7]	Complex background and different lighting condition	Robust to lighting changes
Camshift algorithm for Tracking [9]	Simple background	More rapid results than mean shift algorithm
YCbCr color Model [1]	Different light conditions	Less accurate during dynamic gestures
HSI, YCbCr and morphological operations[13]	Different luminance conditions and complex backgrounds	Provides enhanced performance and Accuracy as compare to the YCbCr and HSI methods used alone
Otsu's Thresholding [5]	Simple background	Assumption of uniform illumination
Tracking by Viola Jones method [14]	Different lighting conditions and Simple background	Accurate and quick learning method for object detection

3.2 Feature Extraction

The most fundamental step in recognition of sign language is Feature Extraction. This step extracts important features after hand tracking and segmentation has been carried out successfully. This step helps in data dimensionality reduction. Feature extraction is very crucial to sign recognition performance. The most significant decisions in the design of gesture recognition system is the selection of appropriate features and feature extraction methods. The feature vectors are obtained in this step. These feature vectors are given as an input to the next step i.e. classification step where these vectors are used to train the classifier. Various feature extraction techniques are Fourier descriptors [1], Hierarchical centroid [2], HOG features [10], PCA [5], Hu Invariant Moment [6], Structural shape descriptors [6], etc.

TABLE-2: Feature Extraction Techniques

Methods	Literature Availability	Remarks
Hu Invariant Moment [6]	Different Background and disjoint shapes	Calculation of seven moment values of on the hand
Principle Component Analysis [5]	Occluded and overlap gestures can be recognized	Not scale invariant and more training needed
Fourier Descriptors [1]	Simple, robust, computationally efficient and immune rotation or Scaling of shape and noise.	Local features can be located
HOG descriptors [10]	Invariant to illuminat change and orientation	Not scale and rotation invariant

3.3 Classification

In order to classify the signs which are given as an input to the system into different classes, a classifier is needed. The output of the feature extraction step i.e. a feature vector is keyed into the classifier which ultimately identifies the sign [3]. Classification step involves two phases: training phase and testing phase. During the training phase, the feature vectors obtained in the previous step are given to the classifier. The testing phase involves the identification of the class equivalent to the sign, when a test

image or video is given as an input and finally the result is produced either as a text or a voice. Most commonly used classifiers are Euclidean distance [1], K Nearest Neighbour (KNN) [2], Artificial Neural Network [3], Support Vector Machines (SVM) [4], Genetic Algorithm [9], Possibility theory [12], Hidden Markov Model [15], etc. The classifier's performance is measured in terms of rate of recognition.

TABLE-3: Recognition Techniques

Methods	Advantages	Disadvantages	Acc %
Support Vector Machine [4]	Accurate results, overfitting and robust to noise	Computationally expensive and runs slow	94.23
Euclidean distance [1]	Invariant to rotation	Time consuming	97.50
Artificial Neural Network [3]	Low computation complexity	Time consuming	91.11
K Nearest Neighbour (KNN) [11]	Robust to noisy data	Need to determine value of parameter	80.00
Hidden Markov model [15]	Allows variant structures to be modelled directly	Expensive in terms both memory and compute time.	93.70

4. RELATED WORK

Purva C. Badhe and Vaishali Kulkarni proposed a system for recognition of ISL gestures from a video [1]. The proposed scheme converts ISL numerals and alphabets into English. They used a combinational algorithm which included canny edge detection, skin color detection with YCbCr, thresholding, etc. for tracking hand movement. They used Fourier descriptors method for feature extraction. The proposed system makes use of template matching for recognition of signs. The database used for implementation was self-created which included a total of 130,000 videos out of which 58,000 videos were tested for checking the performance of the system. The proposed system achieved an accuracy of 97.5%. The complete system was developed in MATLAB by creating a Graphical User Interface (GUI).

Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo proposed a system for automatic recognition of ISL numerals [2]. They used only a simple regular digital camera to obtain the signs. A database containing 5000 signs, 500 images per numeral sign was created by them. Direct pixel value method and hierarchical centroid method were used to extract required features from images of numeric signs. Classification of the signs was done using neural network and kNN classification techniques. The overall recognition rate reported was 97.10%.

The work in [3], presents a method based on Artificial Neural Network. They developed a system for recognition of static gestures in ISL alphabets and numerals automatically. The proposed system translates fingerspelling in ISL to textual form. The gestures included English alphabets (26 letters) and numerals (0-9). They used YCbCr model and also applied filtering and morphological operations for hand segmentation. A novel method based on the distance transformation of the images was projected. Artificial neural network was used to classify the gestures. The system was tested for 36 signs with 15 images of each. The system was implemented using MATLABR2010a. The proposed method had a low computational complexity. A recognition rate of 91.11% was achieved.

A. Singh, S. Arora, P. Shukla and A. Mittal in [4] proposed a system which was capable of classifying ISL static gestures captured under indistinct conditions. The system partitioned gestures into single handed or double handed gestures. Geometric descriptors and HOG features were used to extract features. A database consisting of 260 images captured under simple and complex backgrounds was collected for the experimental purpose. The proposed approach compared KNN and SVM classifiers and it was concluded that SVM was superior to KNN algorithm in terms of accuracy on both geometric and HOG features. SVM achieved the highest accuracy of 94.23%.

A real time recognition of ISL gestures proposed by Shreyashi N. Sawant and M. S. Kumbhar was presented in [5]. They used Otsu algorithm for segmentation purpose. PCA was used to lessen the feature vector for a particular gesture image. A dataset consisting of 260 images, 10 each of the 26 signs was used. The images were captured at a resolution of 380 x 420 pixels. Euclidean distance was calculated between test and train image and gesture having minimum distance was recognized. Recognized gesture was transformed into text and voice format and text was displayed on GUI screen.

K. Dixit and A. S. Jalal in [6] proposed an approach for translating ISL gestures into usual text. Global thresholding algorithm was employed for segmentation purpose. Features were extracted using Hu invariant moment and structural

shape descriptors. Multiclass Support Vector Machine was used as a classifier and maximum likelihood selection was done to recognize the gestures. A dataset consisting of 720 images was used. A recognition rate of 96% was achieved.

Ghotkar A. S., R. Khatal, S. Khupase, S. Asati and M. Hadap in [9] proposed their work which consisted of four modules for recognition of sign languages. CAM-SHIFT method was employed for hand tracking. HSV color model and neural network were used for hand segmentation. For feature extraction, Generic Fourier Descriptor (GFD) method was used. Genetic algorithm was used for gesture recognition. The authors did not mention about the database and no result was reported about their work.

5. CONCLUSION

This paper presents a survey on various hand gesture recognition approaches for Indian Sign Language. An important goal of gesture recognition system is to build an efficient human-machine interaction. In the recent years, many researchers have mainly concentrated on recognition of gesture that has been recorded under static hand gesture. Only few works have been reported for recognition of dynamic hand gesture. The majority of the systems are dependent on the signer. Also, facial expressions are not included in most widely used systems. Developing systems which are capable of recognizing both hand and facial gestures is a key challenge in this area. Comparison between various gesture recognition systems have been presented in this paper. To achieve better results and high accuracy, future investigation in the areas of segmentation techniques, feature extraction methods and classification methods are required to achieve the final objective of human computer interface in the area of sign language recognition for hearing and speech impaired people.

6. REFERENCES

[1] 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.

[2] Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo , "Indian Sign Language Recognition Using Neural Networks and KNN Classifiers", ARPN Journal of Engineering and Applied Sciences, Vol. 9, No. 8, August 2014.

[3] 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.

[4] A. Singh, S. Arora, P. Shukla and A. Mittal, "Indian Sign Language gesture classification as single or double handed gestures," 2015 Third International Conference on Image Information Processing (ICIIP), Waknaghat, 2015, pp. 378-381.

[5] 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.

[6] K. Dixit and A. S. Jalal, "Automatic Indian Sign Language recognition system," Advance Computing Conference (IACC), 2013 IEEE 3rd International, Ghaziabad, 2013, pp. 883-887.

[7] Shreyashi Narayan Sawant. "Sign Language Recognition System to aid Deaf-dumb People Using PCA", International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 5 No. 05, May 2014.

[8] Geetha M, Manjusha U. C, "A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation" International Journal on Computer Science and Engineering (IJCSE), March 2012.

[9] A. S. Ghotkar, R. Khatal, S. Khupase, S. Asati and M. Hadap, "Hand gesture recognition for Indian Sign Language," Computer Communication and Informatics (ICCCI), 2012 International Conference on, Coimbatore, 2012, pp. 1-4.

[10] S. C. Agrawal, A. S. Jalal and C. Bhatnagar, "Recognition of Indian Sign Language using feature fusion," Intelligent Human Computer Interaction (IHCI), 2012 4th International Conference on, Kharagpur, 2012, pp. 1-5.

[11] B. Gupta, P. Shukla and A. Mittal, "K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion," 2016 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, 2016, pp. 1-5.

[12] N. Baranwal, K. Tripathi and G. C. Nandi, "Possibility theory based continuous Indian Sign Language gesture recognition," TENCON 2015 - 2015 IEEE Region 10 Conference, Macao, 2015, pp. 1-5.

[13] Avinash Babu.D, Dipak Kumar Ghosh, Samit Ari; "Color Hand Gesture Segmentation for Images with Complex Background"; International Conference on Circuits, Power and Computing Technologies; IEEE 2013.

[14] L. Yun, Z. Peng, "An Automatic Hand Gesture Recognition System based on viola-Jones Method and SVMs", International workshop on Computer Science and Engineering, IEEE Computer Society,2009,pp.72-76.

[15] Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang; "Hand gesture recognition using a real-time tracking method and hidden Markov models"; Image and Vision Computing; Science Direct 2003.