A Comprehensive Analysis on Sign Language Recognition System

Rajesh George Rajan, M Judith Leo

***Abstract*: Thriving efforts in the area of Sign Language Recognition (SLR) research within the last few decades makes a good interaction between human and computer system. Sign Language is basically a means for dissemination through signing which, utilizes specific sign patterns performed to deliver the meaning with the use of hands, lips and facial expressions to conveniently be able to express the signer's thoughts. The process involves the concurrent association of the shape of hands, the position of the body posture and also the facial expressions. We provide a review of different Automatic Sign Language Recognition system reported in the last few years.**

***Index Terms*: Sign Language Recognition System, Gesture Recognition and Hand Gestures.**

I. INTRODUCTION

The hearing impaired people find it very difficult to express their feelings to the normal people, since the normal people are unaware about the sign language used by the deaf and dumb people. Sign language is the main source of medium for conveying messages, emotions, feelings etc. for deaf and dumb people and hard hearing community. This is also helpful in communication between speaking community and the person who neither speaks nor hears. Nonverbal communication, that conveys messages via hand and body movements, facial expressions that fulfills most of the communication among human beings [1]. A sign language is also termed as a group of feasible forms of gesture communication. Sign language communication involves both manual and non-manual signs [2].

To produce a better communication between the people who neither speaks nor hear and the normal people, the below combination of three channels should be used. Firstly, the hand gesture is the familiarsort of body language is used mostly for conveying messages. The hand gesture is commonly divided into static and dynamic based on the viewable features. The hand position which does not alter during the signing period is known as static hand gestures. It mainly depends on the shape and gyratory angles of the fingers. On other hand, the hand position which alters continuously with respect to time is called as dynamic gestures. It generally has three motion phases: preparation, fondle, and retraction [3]. Secondly, the facial expression is used, where signers show various facial expressions in order to express different emotions, feelings etc. The third channel is the body movements like the head and the upper body movements. Here, the facial and body movements are used to convey the ideas of syntactic and semantic information in the sentence. Sign Language Recognition (SLR) system is the technique to identify a sequence of developed signs and translated it into text or a speech with proper meaning. SLR is a combination of research field including pattern recognition, natural language processing, computer vision and linguistics [4]. SLR systems can be used as a connection between an individual and computer systems. The sign language recognition can be done either in two ways. One is with physical devices like gloves, Kinect etc. and the other is computer vision based. The movements were captured by these sensor devices easily and accurately, because these devices are attached to performer's body. But the system is not user friendly due to the reason that we cannot take these devices to every place. These type of issues is avoided by introducing computer vision algorithms for getting the features. The main components of the system as shown in the Fig.1., and the videos or images are used as the probe data and these probe is used for pre-processing, feature extraction and then it is stored in the database of training. Dimensionality reduction and from that output the video or image is classified as the corresponding word or sentences.

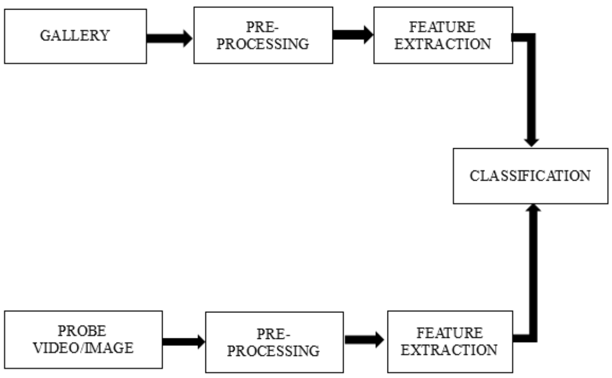

Fig.1 Sign Language Recognition components

The remaining paper is structured as follows. In section 2, we discuss about the different methods used for the sign language recognition and in section 3, we have mentioned the discussion part and finally a conclusion and future perspective.

Revised Manuscript Received on 30 March 2019.
* Correspondence Author
Rajesh George Rajan,* Research Scholar, Hindustan Institute of Technology and Science, Chennai, India.
M Judith Leo, Assistant Professor, Hindustan Institute of Technology and Science, Chennai, India.

II. DIFFERENT ACQUISITION METHODS USED FOR SIGN LANGUAGE RECOGNITION

Different methods are available in the sign language recognition system. The common devices used for the data acquisition are camera, glove, Kinect and leap motion based devices. In this paperwe have classified the sign language recognition methodologies based on the acquisition devices.

2.1 CAMERA BASED ACQUISITION

The camera based or vision based approach, acquires the movements of the performer's hand using a camera. Some authors [6], [7], [11], [12] were used coloured or painted glove in order to easily identifying the position of the fingers. The camera based SLRS categorize the gestures from a 2D image with the help of machine learning and image processing techniques. This type of acquisition, the users feel free while signing and the cost of the system is very low and be more powerful because of their flexibility, portability, user friendly [19].Several authors have used camera based data acquisition system, which is discussed below.

Al-Jarrah et.al.[5], developed an Adaptive Neuro-Fuzzy Inference Systems (ANFIS)to recognize the 30 manual words or gestures from the Arabic sign language. For each gesture, 60 samples were taken by 60 different subjects. Apart from the other hardware the author used camera for image acquisition and it is convenient for the common people. The major parts of the recognition framework include image or video collection, pre-processing, segmentation, feature extraction and classification. In the feature extraction method, they used the border information, the center of area, and the direction of the gesture to elicit a feature vector for the gesture. The system achieves a recognition accuracy of 93.55%.

E. E. Hemayed et.al.[8] developed a classification framework based on an Arabic sign language alphabet which transforms signs into voice. The system mainly focuses on passive and simple changeable gestures. The dataset consists of 150 signs and gesture images. The colour images of the gesture inputs are acquired with the help of web camera. To abstract the skin blobs, the YCbCr space was used. To separate the hand gesture, the Prewitt edge detector is used. K-Nearest Neighbour Algorithm (KNN) with Principal Component Analysis (PCA)was used in the recognition stage. The system obtained a classification rate of about 97%.

Table.1 Summary of Camera Based Acquisition

Ref & Year	Dataset/ Language level /Number of subjects	Classifier	Recognition rate
[13],2013	8 hand gestures,/N.A	SVM+ANN	Not-Mentioned
[11],2012	300 Arabic Sign Language gestures/ N.A	HMM	95
[16],2011	26 ISL gestures/ N.A	SVM	91.3
[8],2011	150 gestures/ N.A	PCA+KNN	97
[9],2010	30 Arabic alphabet gestures/N.A	MLP+MDC	91.3
[6],2008	30 gestures, 900 samples/ 2 subjects	Fully Recurrent network	95.1
[7],2005	42 gestures of ArSL,2323 samples/N.A	Polynomial classifier	93.4
[14],2005	18 signs/6 subjects	HMM	87.8
[17],2002	65 JSL words/ N.A	HMM	Not-Mentioned
[18],2003	20 gestures/ N.A	HMM	90
[5],2001	30 Arabic sign words / 60 subjects	ANFIS	93.55
[15],1998	500 ASL sentences having 40 words/N.A	HMM	92

In [6], M. Maraqa et.al., proposed an Arabic alphabet classification framework with fully recurrent neural networks. The dataset consists of 900 specimens of Arabic sign language with single handed images. It consists of 30 gestures performed by 2 signers. A digital camera and coloured gloves [7], [11] were also used for the image acquisition in their experiments. Border information, the center of area, and the direction ofthe gesture is used to obtain a feature vector for the gesture. In this system two neural network architectures were used; an Elman recurrent network and a fully recurrent neural network which has full feedback loops. The Elman network achieves a recognition rate of 89.7% although a fully recurrent network enhanced the efficiency of 95.1%.

K. Assaleh et.al. [7], proposed a framework to recognize 42 gestures from Arabic alphabet signs using polynomial classifier. A coloured glove having 6 different colours is used in this system. This glove is five for fingertips and one for the wrist region. The recognition is obtained with the help of ANFIS network system. The features were different spatial measures such as angles and lengths. 93.4% accuracy rate was obtained.

Mohandes et.al.[11], used isolated images of Arabic signs to recognize the system with Hidden Markov Model (HMM). Skin colour detection approaches [14], [15], [16], [17], [18] were used in different methods to find out the segmentation of hands. A hybrid Gaussian skin colour model was used to find out the signers face and with centroid to the face region, the movements of the hands are to be detected. The issues of hand region segmentation is alleviated by wearing two coloured gloves in both the hands. The dataset consists of 500 samples of 300 signs and the system accomplishing a recognition rate of 95%. R. Wang et al[12] proposed a hand tracking approach in real time with one camera and a cloth glove that is printed with different colour pattern. They designed the pattern to make easy the pose estimation dilemma, so that it is easy to apply a nearest-neighbor approach (k-NN) to track the hands. The author obtained a set of 18,000 finger configurations with the help of Cyber glove II motion capture system. They introduced a new desktop virtual reality application with the help of the hand motion. In [9], El-Bendary et.al., proposed a recognition framework based on 30 Arabic manual alphabet. This system uses a video of signs and that video is translated in to corresponding text.

The outline of hand is extracted from each frame and each best-frame will go through edge-detection stage, and then finally followed by feature-vector-creation stage. The obtained features are rotation, scale, and translation invariant. For the classification stage, a multilayer perceptron (MLP) network and a minimum distance classifier (MDC) were used. This system achieved 91.3% accuracy. Y. Quan et.al., [10], proposed a Chinese sign language classification structure which is relatedto obtaining the features of video, while the signer is performing the letters of Chinese language. In this they used five image descriptors to define different visual and geometrical properties like the number of interest points, colour histogram,48 dimensional Gabor wavelet,7Hu moments, 128 Fourier descriptors, and their SIFT features. SVM classification method is used to recognise the letters and an average classification rate of 95.55% was obtained. J. Zieren et.al. [14], proposed a framework with video clips as the input to classify the sign language. Through adapting the colour of the skin, the hand segmentation and mobility tracking was achieved. Using HMM, the spatial features from the segmented images of 18 sign gestures by 6 subjects was extracted with an average accuracy of 87.8%. They used a directory of British Sign Language video clips. Initially they recognise user-dependent system having 232 isolated signs with an efficiency of 99.3% and then user independent system with 221 signs having a recognition rate of 44.1%. T. Starner et.al. [15] proposed a classification system of continuous American sign language (ASL) sentence with the help of a single camera. To find out the hand segmentation and hand-blobs,the skin colour was used where the feature vector of 16 dimension was computed. After test with a dataset of 500 sentences and a 40-word lexicon, the recognition rate of the experiment was 92%.

In J. Rekha et.al. [16],presented a video related SLR framework for alphabets from Indian Sign Language (ISL). For hand segmentation and detection, the author adopts a skin tone modelling. The texture,hand shape and figure count features were measured using Wavelet Packet Decomposition (WPD-2),Principle Curvature Based Region (PCBR) and Convexity defects algorithm respectively. The classification process was performed on 26 ISL alphabets with k-NN, Support Vector Machine (SVM) and Dynamic Time Warping (DTW) classifiers and the best recognition rate was obtained by using multiclass SVM classifier with a rate of 91.3%. In [17],N Tanibata et.al.,discussed aJapanese Sign Language (JSL) classification framework using skin-color based segmentation. Using skin colour and elbows, they followed the face and hands for every frame. The Overlapped hand and face were tracked by paring the texture template of the previous regions of the hands and face. The 65 JSLwords whose face and hands were classified accurately with HMM. Chen et.al. [18] developed a dynamic continuous gesture classification method with background modelling, skin colour, edge based features and motion. The Fourier descriptor (FD) and the temporal features are characterized by spatial features and the motion analysis respectively. The approach used for data acquiring is the camera and HMM for the classification of the various gestures.With the help of motion, skin colour and edge data, the position of hand was detected. Here they only use single hand to create hand gestures. Their system is used to classifyvarious20 gestures, and the accuracy is greater than 90%. The above table 1 show the summary of the camera based acquisition.

2.2 GLOVE BASED ACQUISITION

To recognize the hand gestures, an electro mechanical device is placed inside a glove. So the deaf and dumb people are recommended to wear a glove in order to gather information. This glove is linked to some sensors and these sensors acquire the gesture information related to shape, movements, locations etc., of the hands [13].The below observations are based on the glove based data acquisition.

M. A. Mohandes [19], proposed a recognition system based on Arabic language. Here there were both the hands are used for the data acquisition. In this system the activities are done with the help of two cyber gloves and two hand tracking devices like Flock of Birds (FOB). There was two subjects to perform 100 double handed sign from which 20 samples was taken in to the database. Second order statistics from sub-frames of the signs were used as features. The dimensionality of the feature vector was diminishedwith the help of Principal Component Analysis method. Having SVM with 100 signs, the system recorded anaccuracy of 99.6%.

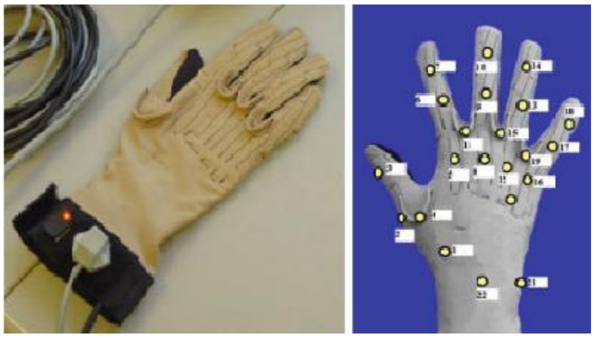

Fig.2 Cyber Glove

X. Zhan et.al.[20], discussed a system for Chinese sign language CSL) classification systemhavingcomposite of 3D accelerometer sensors and multi-channel electromyography (EMG). By using the EMG signals the gesture divisions were detected.Using combined HMM and Decision Tree method,this system has been classified 72 sign language words. In addition to this, 40 CSLsentences was implemented and recognised to figure out their framework for sequential sign language. The accuracy of overall words and sentences are 93.1% and 73.5% respectively. In [21], N. Tubaiz et.al., obtained 40 Arabic sign language sentences for classification with a camera and two DG5-VH data gloves. For a dataset having sequential in nature, the authors proposed the Modified K-Nearest Neighbors classifier (M-KNN) to obtain the recognition rate of 98.9%sentence-based sign language. N. M. Kakoty et.al. [22], collected 8 Indian sign alphabets,26 American sign alphabets, and 0-9 digits using data glove. In this system kernel support vector machine with 10-fold cross validation was used for recognition. The average recognition rate of 96.7% was achieved for the ISL, ASL and sign numbers. They were translating the corresponding sign language in to speech by using label matching techniques.

Table.2 Summary of Glove Based Acquisition

Ref & Year	Dataset/ Language level /Number of subjects	Classifier	Recognition rate
[22],2018	ASL alphabets, 0-9 numbers, 8 alphabet of ISL/N.A	Kernel-SVM	96.7%
[25],2018	27 ASL alphabets/N.A	Embedded-SVM	98.2
[26],2017	25 ISL single handed dynamic sign words/N.A	Coupled HMM	90.80
[27],2017	50 ISL dynamic words/N.A	HMM & BLSTM-NN separately	97.85 & 94.55
[24],2016	150 words/N.A	Multistream HMM	94.31% & 87.02% for user dependent and independent
[21],2015	40 ArSL sentences/N.A	Modified K-Nearest Neighbors classifier (M-KNN)	98.9%.
[19],2013	20 samples of ArSL /2 subjects	SVM	99.6
[28],2011	50 isolated words of ASL gestures/N.A	ANN	90
[20],2011	72 sign words of CSL / 40 CSL sentences/N.A	Combined HMM & Decision tree based classifier	93.1 &72.5 for overall words & sentences

X. Yanget.al.[24], proposed a Chinese Sign Language sub word classification system with surface electromyography (sEMG), accelerometer (ACC), and gyroscope (GYRO) sensor. They used 150 regularly used nouns and verbs and it include single and double handed signs and conducted using user independent and user specific manner. Optimized tree-structure based classification i.e., Multistream HMM was used for the classification purpose. The recognition rate for user-dependent and user-independent was 94.31% and 87.02% respectively.

B. G. Lee et.al.[25], developed a smart wearable hand device sign language interpretation system. This framework consists of five flex-sensors, two pressure sensors, and a three-axis inertial motion sensor and it is used to recognize the real time 27 ASL alphabets. This system was developed in android based mobile, and the translation of text-to-speech of the alphabets is the one of the functionality. A built-in embedded SVM was used for the recognition. With the help of the pressure sensor placed on the middle finger, the classification rate is increased to 98.2%.

P. Kumar et.al. [26], presented a recognition framework for the Indian sign language using multi-sensor framework. The pre-processing and feature extraction is in connection with the leap motion and Kinect sensors. Fingertip location and direction methods were used in the feature extraction. The methods like early and late fusion techniques are used for the gesture classification. This ISL dataset consists of 25 single handed dynamic words. The classification accuracy is achieved as 90.80% with Coupled Hidden Markov Model. In [27], P. Kumar et.al., again clearly discussed a framework related to leap motion and Microsoft Kinect sensors. In this system they used both hands for gesture recognition. This system used 50 ISL dynamic words for recognition. Classification is carry out individually by HMM and Bidirectional Long Short-Term Memory Neural Network (BLSTM-NN) based sequential classifier and accuracies of 97.85% and 94.55% have been recorded.

Cemil Oz et.al.[28], developed a system that translate ASL sign words to English by using Artificial Neural Networks. The proposed system used Cyber glove and Flock of Birds trackers in order to get the features of the gestures. Artificial neural network is used to classify the features of gestures. The dataset having 50 isolated words of ASL and the accuracy of the system is 90%. Table 2 shows the summary of the glove based acquisition.

2.3 KINECT BASED ACQUISITION

Microsoft Kinect is the hardware system used to acquire the 3D level of information or depth and skeletal information. With the 3D positions of the body joints, we can easily track the movements of the body very accurately. For the 3D estimation of the person or objects in the space, Microsoft Kinect provides a good solution.The below discussions is based on the Microsoft Kinect based data acquisition.

In [29], N. Pugeault et.al., created a finger spelling ASL system in continuous environment with the help of Kinect. Based on the Gabor filtering they extracted the hand shape features and Random forest is used to classify the letters of the ASL having recognition rate of 75%. The authors took around 48000 samples of the signs. i.e. 500 for each sample. This system also included the dynamic letters from the ASL alphabets.

Fig.3 Microsoft Kinect Sensor for Xbox 360

C. Zhang et.al. [30], presented a visual recognition framework based on individual ASL signs. This framework proposed a multimodality way to find out the recognition of ASL. So it includes hand gestures, facial expressions and body movements. ASL videos are acquired by Kinect sensor and a total of 61 video sequences were used. Linear SVM is used as the classifier and the classification rate of the corresponding system was 36.07%. R. Saini et.al. [31], implements a Human Activity Recognition (HAR) framework with the help of Microsoft Kinect, that is flexible and continuous.

The system used a huge amount of dataset, around 1110 sequences encompassed of 24 isolated operations. The system records the different exercises in to a group like activities during sitting and standing and then given to the BLSTM-NN network for classification. The classification rate of the corresponding system was 68.9%.

Table.3 Summary of Kinect Based Acquisition

Ref & Year	Dataset/ Language level /Number of subjects	Classifier	Recognition rate
[31],2018	1110 sequences of 24 isolated signs/N.A	BLSTM-NN	68.9
[33],2016	25 Taiwanese words/N.A	SVM-HMM	85.14
[30],2016	61 ASL video sequences/ 5 subjects	Linear SVM	36.07
[34],2015	25 words/N.A	3D-CNN	94.2
[35],2015	35 words of ISL/N.A	SVM	86.16
[36],2012	24 words of ASL/N.A	Randomized Decision Forest (RDF)	84.3
[29],2011	48000 samples of ASL/N.A	Random forest	75

H. Wanget.al. [32], present a Kinect based SLR framework using RGB-D data. The proposed system was based on sparse observation i.e. the dynamic motion of hand trajectory, where the speed of the hand should be in the minimum. Hand motions are represented using 3D hand and hand postures were represented using HOG features. The hand posture is correctly described by HOG and grouping through K-means clustering. They created two types of datasets namely daily SL dataset and large vocabulary SL dataset having 370 and 1000 daily and isolated signs respectively.

G. C. Lee et.al. [33], discussed a Taiwanese gesture based communication acknowledgment framework utilizing Microsoft Kinect. The position, direction and shape of the hands are the main features extracted. 25 Taiwanese words were used for the experiments. The direction of hand movements is determined by HMM and SVM is used to classify the hand shapes. The accuracy of the system was 85.14 %. Huang. J, et. al., [34], developed a3-D convolutional neural network model for recognition. Due to the large variations in hand actions, the classification of the sign language is difficult. So that the tracking and segmentation of hands with the help of Kinect sensor. The dataset has 25 vocabularies that are commonly used in daily life. This method has a recognition rate of 84.3%.

K. Mehrotra et.al. [35], developed an Indian SLR system with 3D skeleton points with the help of Kinect sensor. This system obtained 37 non-continuous sign gestures of ISL. These gestures having skeleton joints of the upper body in order to extracting angular and distance based features. The SVM classifier is used for recognize the gestures having a rate of 86.16%. C. Keskin et.al. [36], developed a Kinect based ASL hand gesture recognition and position estimation system. The system having ASL alphabets with 24 static sign gestures. In this system each pixel was classified by Randomized Decision Forest (RDF) and the majority voting based scheme was used to find out the final class label. This method has a recognition rate of 84.3%.The below table. III show the summary of the Kinect based acquisition.

2.4 LEAP MOTION SENSOR BASED ACQUISITION

Leap motion is a software and hardware which develops 3D motion control technology which supports hand and finger motion or action. Several researchers have used this Leap motion for establishing SLR, gaming, robotics and other HCI systems [27].The below observations are based on the leap motion sensor based data acquisition

L. Quesada et.al.[37],proposed a recognition framework with Leap Motion and Intel RealSense, the devices utilized for hand tracking. Classification is done using Support Vector Machine over 50 subjects.

R. Kadry et.al. [38], developed a classification system on ASL using leap motion sensor. They collected a dataset of 3600 images for ASL having 24 static letters and 10 numbers, and 2 images for dynamic letters 'Z' and 'J' from different signers. Feature extraction using discrete cosine transform (DCT) for dynamic letters and ANN is used as a classifier to give an average success rate of 87%. This system is also supported with; a robotic arm that is built using Arduino Uno Micro-Controller and two servos motors to track the hand. Weighted average success rate of 83.11% accuracy is achieved in this approach.

L. E. Potter et.al. [39], proposed a SLR system with Leap motion sensor. The system classifies the basic signs of Australian Sign Language (AuSL) in order to evaluate the recognition. For classification, ANN was used. The proposed system was not able to classify difficult signs.

M. Mohandes et.al.[40],proposed a system using the Leap Motion Controller that classifies the twenty-eight Arabic sign language. The classification is done by using both Naive Bayes Classifier (NBC) and the Multilayer Perception (MLP) and trained by the back-propagation algorithm. It was found that MLP achieves highest recognition rate of 99.1%.

Table.4 Summary of Leap Motion Based Acquisition

Ref & Year	Dataset/ Language level	Classifier	Recognition rate
[38],2017	3600 images in ASL	ANN	87
[41],2015	22 letters of ASL	Genetic algorithm	82.71
[40],2014	28 alphabets of ArSL(2800 frames of data)	Naive Bayes & MLP	99.1
[42],2014	50 dynamic signs of ArSL sign	MLP-NN & Naive Bayes	88

M.Funasakaet.al. [41] developed a finger spelling language classification method using Leap Motion Controller. The dataset consists of only static 22 letters of ASL. Putting hands and fingers over a Leap Motion controller, the finger spelling recognition is performed. Leap Motion has skeletal tracking that recognizes the framework of fingers.

The system used decision tree that use 16 conditions related to finger and hands. The decision tree is automatically generated by a Genetic algorithm to obtain quasi-optimal solutions. The system has a recognition rate 82.71% using decision tree based approach.

A. Elonset.al. [42], proposed an Arabic Sign Language (ArSL) framework having 50 different dynamic signs with leap motion sensor. The classifiers in this system are Multilayer Perception Neural Network (MLP-NN) and Naive Bayes and recognition rate of 88%. The below table 4 show the summary of the Leap motion based acquisition.

III. RESULT AND DISCUSSION

The sign language recognition is highly effective and productive system which helps to tackle the obstacle between the hearing impaired and the normal people. Sign language recognition has become one of the most common studies that have been carried out now in the field of human computer interaction widely gaining global attention. However, SLR has got many limitations and in order to reach the desired goal.

The different barriers subjected to acquire data from videos such as the number of cameras used, placement of the camera, environmental conditions like luminous sensitivity, background condition, colour of signers and their textile etc. Another issue faced in the sign language classification is the segmentation. The segmentation is obtained by two main techniques. One segmentation is based on different external aids like data and colour gloves and the other is skin colour based segmentation. Some authors [6], [7], [11], [12] were used coloured or painted glove and the others used skin colour [14], [15], [16], [17], [18] based segmentation. Lighting variation, background colour, difference in skin colour etc. is also a problem in segmentation.

Another major challenge of sign language classification is extracting features and the classification accuracy. The multimodal way of recognizing sign language is difficult. The inclusion of the different modalities in the sign language is the proper way to convey a sign. Normally it was avoided by the researchers due to difficulty and time complexity. However,[17], [43], [44] includes hands along with faces. A sign is also affected by the current and preceding sign, i.e., co articulation.

Another issue is tracking of the sign. The time and space differs in each performer and their signing velocity differs every time. The repeated signs of one performer can even cause minute changes in the speed and position of the hands. The combination of the manual and non-manual sign is also an issue. The main challenges are usage of both manual and non-manual signs simultaneously. The static signs are easy to classify.

There is also a challenge in the geographical variation in different sign language. No standardized dataset is available in country like India, which affect the research related to the sign language recognition.

Another drawback of the recognition system is the usage of various hardware devices like sensors, gloves etc. in this process. Because they have high costs and it is difficult to carry. So the major challenge is to integrate the system with good recognition using computer vision techniques.

IV. CONCLUSION

The sign language is the tongue of the people who neither speak nor hear. Without the sign language they do not exist in the world. The human-machine interaction is developed through the gesture recognition system. In the previous years, most of the researchers have done their research in static hand gesture recognition. Some works have been reported for recognition of dynamic hand gesture. 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. In this paper we have discussed different sign language recognition approaches using different acquisition methods. By using the different data acquisition methods like sensor based gloves, Kinect, leap motion controller etc. and these devices having good recognition rate, but the computation complexity and cost is too high and having difficulty to maintain in the public places. So in order to avoid the severe problem, the camera based approaches is more beneficial. With the help of web camera, we have acquired the input data easily and processed. 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 deaf and dumb people.

REFERENCES

1. K. Hogan, R. Stubbs,“ Can’t get Through 8 Barriers to Communication”, Pelican Publishing Company, Grenta, LA, 2003.
2. Mitra, S., Acharya, T., “Gesture recognition: A survey”. IEEE Trans. Systems Man Cybernet. Part C Appl. Rev. 37 (3), (2007).
3. A. Kendon,“Current issues in the study of gesture, in: The Biological Foundation of Gestures”: Motor and Semiotic Aspects, Psychology Press, 1986, pp. 23–47.
4. O. Aran, I. Ari, L. Akarun, B. Sankur , A. Benoit , A. Caplier, P. Campr, A. H. Carrillo, and F. Xavier Fanard, “SignTutor: An Interactive System for Sign Language Tutoring,” IEEE feature article, pp. 81-93, 2009.
5. O. Al-Jarrah and A. Halawani, “Recognition of gestures in Arabic sign language using neuro-fuzzy systems,” Artificial Intelligence, vol. 133, no. 1, pp. 117–138, 2001.
6. M. Maraqa and R. Abu-Zaiter, “Recognition of Arabic sign language (ARSL) using recurrent neural networks,” in Applications of Digital Information and Web Technologies, 2008. ICADIWT 2008. First International Conference on the. IEEE, 2008, pp. 478–481.
7. K. Assaleh and M. Al-Rousan, “Recognition of Arabic sign language alphabet using polynomial classifiers,” EURASIP Journal on Applied Signal Processing, vol. 2005, pp. 2136–2145, 2005.
8. E. E. Hemayed and A. S. Hassanien, “Edge-based recognizer for Arabic sign language alphabet (ars2v-arabic sign to voice),” in Computer Engineering Conference (ICENCO), 2010 International. IEEE, 2010, pp. 121–127.
9. N. El-Bendary, H. M. Zawbaa, M. S. Daoud, K. Nakamatsu et al.,“Arslat: Arabic sign language alphabets translator,” in Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on. IEEE, 2010, pp. 590–595.
10. Y. Quan, “Chinese sign language recognition based on video sequence appearance modeling,” in Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on. IEEE, 2010, pp. 1537–1542.
11. M. Mohandes, M. Deriche, U. Johar, and S. Ilyas, “A signer-independent Arabic sign language recognition system using face detection, geometric features, and a hidden markov model,” Computers & Electrical Engineering, vol. 38, no. 2, pp. 422–433, 2012.
12. R. Y. Wang, J. Popovi´c, “Real-time hand-tracking with a color glove”, ACM Transactions on Graphics 28 (3) (2009) 63.
13. R. Alzohairi, R. Alghonaim, W. Alshehri, S. Aloqeely, M. Alzaidan, O.Bchir, “Image based arabic sign language recognition”International Journal of Advanced Computer Science and Applications, Vol. 9, No. 3, 2018.
14. J. Zieren, K.-F. Kraiss, “Robust person-independent visual sign language recognition,” Conference on Pattern Recognition and Image Analysis, 2005, pp. 520–528.
15. T. Starner, J. Weaver, A. Pentland, “Real-time American sign language recognition using desk and wearable computer based video”, IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (12) (1998) 1371–1375.
16. J. Rekha, J. Bhattacharya, S. Majumder, “Shape, texture and local movement hand gesture features for Indian sign language recognition”,3rd International Conference on Trendz in Information Sciences & Computing, 2011, pp.
30–35.

17. N. Tanibata, N. Shimada, Y. Shirai, "Extraction of hand features for recognition of sign language words", International conference on vision interface, 2002, pp. 391–398.
18. F.-S. Chen, C.-M. Fu, C.-L. Huang, Hand gesture recognition using a real-time tracking method and hidden markov models, Image and vision computing 21 (8) (2003) 745–758.
19. M. A. Mohandes, "Recognition of two-handed Arabic signs using the cyber glove," Arabian Journal for Science and Engineering, vol. 38, no. 3, pp. 669–677, 2013.
20. X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, J. Yang, "A framework for hand gesture recognition based on accelerometer and EMG sensors", IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 41 (6) (2011) 1064–1076.
21. N. Tubaiz, T. Shanableh, K. Assaleh, "Glove-based continuous Arabic sign language recognition in user-dependent mode", IEEE Transactions on Human-Machine Systems 45 (4) (2015) 526–533.
22. N. M. Kakoty and M. D. Sharma, "Recognition of Sign Language Alphabets and Numbers based on Hand Kinematics using A Data Glove," Procedia Comput. Sci., vol. 133, pp. 55–62, 2018.
23. Z.Zhang. "Microsoft Kinect Sensor and Its Effect". *IEEE multimedia*,19(2):4-10,2012.
24. X. Yang, X. Chen, X. Cao, S. Wei, and X. Zhang, "Chinese Sign Language Recognition Based on An Optimized Tree-structure Framework," IEEE J. Biomed. Heal. Informatics, vol. PP, no. 99, p. 1, 2016.
25. B. G. Lee and S. M. Lee, "Smart Wearable Hand Device for Sign Language Interpretation System with Sensors Fusion," IEEE Sens. J., vol. 18, no. 3, pp. 1224–1232, 2018.
26. P. Kumar, H. Gauba, P. P. Roy, and D. P. Dogra, "Coupled HMM-based multi-sensor data fusion for sign language recognition," *Pattern Recognit. Lett.*, vol. 86, pp. 1–8, 2017.
27. Pradeep Kumar, Himaanshu Gauba, Partha Pratim Roy, Debi Prosad Dogra, "A Multimodal Framework for Sensor based Sign Language Recognition", Neuro computing (2017), doi: 10.1016/j.neucom.2016.08.132.
28. Cemil Oz, Ming. C. Leu, "American Sign Language word recognition with a sensory glove using artificial neural networks."Engineering Applications of Artificial Intelligence 24 (2011) pp. 1204-1213.
29. N. Pugeault, R. Bowden, "Spelling it out: Real-time ASL fingerspelling recognition", International Conference on Computer Vision Workshops, 2011, pp. 1114–1119.
30. C. Zhang, Y. Tian, and M. Huenerfauth, "Multi-modality American Sign Language recognition," 2016 IEEE Int. Conf. Image Process., pp. 2881–2885, 2016.
31. R. Saini, P. Kumar, P. P. Roy, and D. P. Dogra, "A novel framework of continuous human-activity recognition using Kinect," Neurocomputing, 2018.
32. H. Wang, X. Chai, X. Chen, "Sparse observation (SO) alignment for sign language recognition", Neurocomputing 175 (2016) 674–685.
33. G. C. Lee, F. H. Yeh, and Y. H. Hsiao, "Kinect-based Taiwanese sign-language recognition system," Multimed. Tools Appl., vol. 75, no. 1, pp. 261–279, 2016.
34. Huang, J.; Zhou, W.; Li, H.; and Li, W. "Sign language recognition using 3d convolutional neural networks", IEEE International Conference on Multimedia and Expo, 1–6, 2015.
35. K. Mehrotra, A. Godbole, S. Belhe, "Indian sign language recognition using Kinect sensor", International Conference Image Analysis and Recognition, 2015, pp. 528–535.
36. C. Keskin, F. Kırac¸, Y. E. Kara, L. Akarun, "Hand pose estimation and hand shape classification using multi-layered randomized decision forests", European Conference on Computer Vision, 2012, pp. 852–863.
37. L. Quesada, G. López, and L. Guerrero, "Automatic recognition of the American sign language fingerspelling alphabet to assist people living with speech or hearing impairments," *J.* Ambient Intell. Humaniz. Comput., vol. 8, no. 4, pp. 625–635, 2017.
38. R. A. Kadry and A. Birry, "ASL Recognition Using Leap Motion and Hand Tracking Mechanism," Int. J. Adv. Electron. Comput. Sci., vol. 4, no. 9, pp. 2393–2835, 2017.
39. L. E. Potter, J. Araullo, L. Carter, "The leap motion controller: a view on sign language",25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, 2013, pp. 175–178.
40. M. Mohandes, S. Aliyu, and M. Deriche, "Arabic Sign Language Recognition using the Leap Motion Controller". June, 2014.
41. M.Funasaka, Y.Ishikawa, M.Takata and K.Joe," Sign Language Recognition using Leap Motion Controller",2015, pp. 263-269.
42. A. Elons, M. Ahmed, H. Shedid, M. Tolba, "Arabic sign language recognition using leap motion sensor," 9th International Conference on Computer Engineering & Systems, 2014, pp. 368–373.
43. N. B. Ibrahim, M. M. Selim, and H. H. Zayed, "An Automatic Arabic Sign Language Recognition System (ArSLRS)," *J. King Saud Univ. - Comput. Inf. Sci.*, 2017.
44. P. Kumar, P. P. Roy, and D. P. Dogra, "Independent Bayesian classifier combination based sign language recognition using facial expression," *Inf. Sci. (Ny).*, vol. 428, pp. 30–48, 2018.