A comparative Study of Various Techniques and Outcomes of Recognizing American Sign Language: A Review

Shivashankara S[1], Srinath S[2]
[1](Research Scholar, Department of Computer Science & Engineering, Sri Jayachamarajendra College of Engineering, Mysuru, India
shivashankar.research@gmail.com)
[2] (Department of Computer Science & Engineering, Sri Jayachamarajendra College of Engineering, Mysuru, India
srinath@sjce.ac.in)

ABSTRACT

The American Sign Language Recognition (ASLR) process is an evolution in which computer automatically understands the ASL gestures and interprets them into their equivalent human readable text. Several researchers have presented various techniques of different recognition approaches to recognize static and dynamic gestures of American Sign Language. In this paper a sincere effort has been made to highlight the various research works carried out and comparative analysis of those works in recognizing American Sign Language. An attempt has also been made to present sign language recognition approaches and techniques, constraints in gesture recognition, process. Also this paper presents the comparative study and graphical depiction of various techniques of related works carriedout by various researchers.

***Keywords** - America Sign Language, Gesture Recognition, Hand Gesture Recognition (HGR), Sensor Based, Sign Language Recognition (SLR), Vision Based*

1. INTRODUCTION

Sign Language (SL) is a mode of interaction for the hard of hearing people, can converse with the society by distinctive signs and gesticulations. Several countries in the universe have their specific form of SL such as American Sign Language (ASL), Indian Sign Language (ISL), Japanese Sign Language (JSL), Russian Sign Language (RSL), Chinese Sign Language (CSL), and many more.

ASL is a complete and intricate language that uses signs made by moving hands aggregate with body postures and expressions of the face. ASL a core language of many North Americans who are of hearing impaired and also in several nations those who do not have their native SL [1]. As ASL is seen as well-defined and genuine language, it has numerous variations, like other languages do, such as Spanish and French. ASL is a notable form of communication and advantageous to a massive portion of the population. Its basis, current conditions, prospect hopes, and overall impact are quite amazing and eye-opening.

American Manual Alphabet includes a set of 26 signs in ASL, which can be used to spell the words from English language [2]. These signs make use of 19 hand shapes of ASL. The 'P' and 'K' signs uses the same hand shape but there orientations are different. ASL also possesses a set of 10 signs for numeric. Here, there is no native ASL equivalents for proper nouns and technical terms. Fig 1 shows the set of 26 signs of English alphabets and 10 signs of numbers. ASL consists of intricate grammatical structure that desires to be understood. Instead of struggling with perfecting one's accent, however, the challenge of ASL is to integrate the usages of facial expressions and body language.

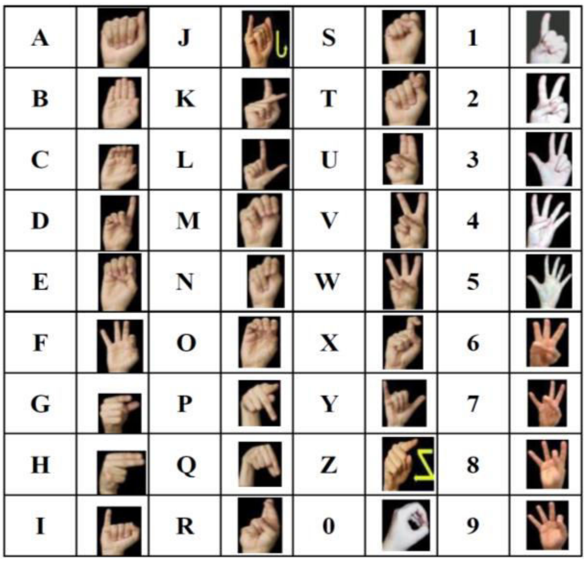

Fig 1. ASL Alphabets and Numbers

The Rest of this review paper is organized as follows. Section 2 describes the Process of SLR; Section 3 defines the various Constraints in Real Time Gesture Recognition. Section 4 describes the distinctive Approaches of SLR; Section 5 explains the various Vision Based SLR techniques and their related works. Finally, Section 6 describe the conclusion.

2. SIGN LANGUAGE RECOGNITION PROCESS

Fig 2 depicts the sequence of steps followed in SLR process.

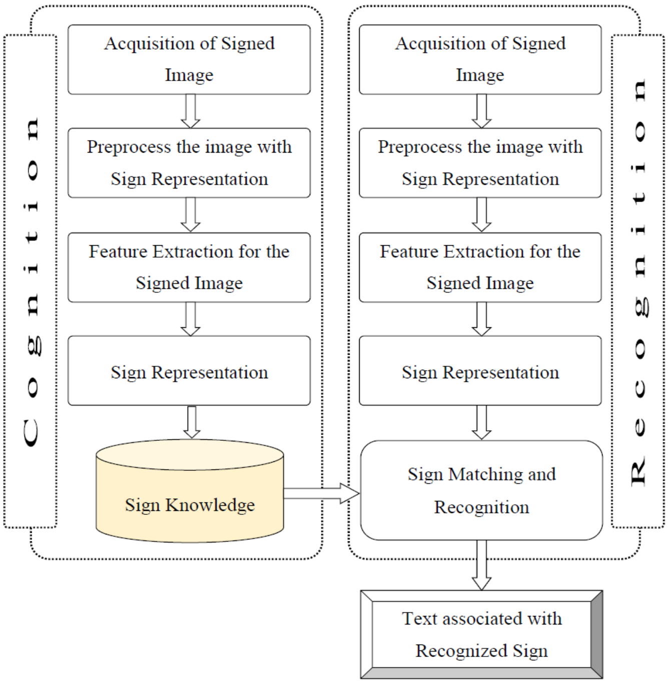

Fig 2. Sign Language Recognition Process

In Cognition process of figure 2, initially signed image is captured from the digital camera and then perform the preprocess operation such as removal of noise, eliminate the blurriness, pixel brightness transformation, geometric transformation and restoration of an image. Later extract the required features from the preprocessed sign image for representing the sign. Once the Sign representation is carried out then all the extracted features of signed image are stored for further utilization in recognition process.

In recognition process, the first four steps are similar as in the cognition process. Sign Matching and Recognition process will be carried out by comparing the stored extracted features with the extracted features of signed image in recognition process. Once the extracted features of signed image are matched with the stored extracted features, then the text associated with the recognized sign will be displayed.

3. CONSTRAINTS IN REAL TIME GESTURE RECOGNITION

Real time gesture is recorded into video using webcam. This video is analyzed for recognition of the gesture. First the video is divided into frames and then the frames are analyzed. The real-time gestures are recognized by considering all the frames at a time and relation between consecutive frames as well. Following are the some of the major constraints in real-time gesture recognition [3].

3.1 Non-uniform, Cluttered and Dynamic Background

The limitation of non-uniform, cluttered and dynamic background can be solved using skin region detection technique. It detects the hand, face, and only skin colored object motion is considered for recognition. Suppose if there are more than one moving objects having skin color then the object of interest considered would be the object having more extent of motion or the object having biggest blob. Different techniques can be used to increase the efficiency of gesture recognition algorithm using techniques called background subtraction.

3.2 Light (Dark or Bright) Illumination Conditions

In vision based approaches different results can be obtained in different illumination conditions, which fails the recognition system. This problem can be solved by adopting more efficient color model against illumination conditions such as YCbCr or HSV instead of RGB model. As compared to models mentioned above RGB is very sensitive to illumination conditions. This leads improvement in the performance but do not completely resolves the illumination problem.

3.3 Image Orientation and Scaling Defects

Image Orientation and scaling defects arises due to the different signers, distance of capturing the gesture of same object. Using SIFT algorithm this defect can be effectively eliminated.

3.4 Processing Speed

The computational time increases due to the limitations on speed of the processor and increased complication of the system. Low density and moderately efficient algorithms could be used resolve this problem.

3.5 Some other constraints

Following are the some of the other constraints of real time gesture recognition [15]:

a) While performing gesture, there may be an Occlusions problem.
b) Position of signer may differ in front of camera while performing gesture.
c) Loss of depth information in working with 2D camera.
d) There may be a change in position and speed with same signer or signer to signer as each gesture differ in space and time.
e) Co-articulation (link between foregoing and succeeding gesture) problem in continuous gesture.
f) Wearing sensor glove always to perform the gesture by the signer is major constraint.

4. SIGN LANGUAGE RECOGNITION APPROACHES

SLR is a significant application of gesture recognition. The following Fig 3 depicts the SLR approaches.

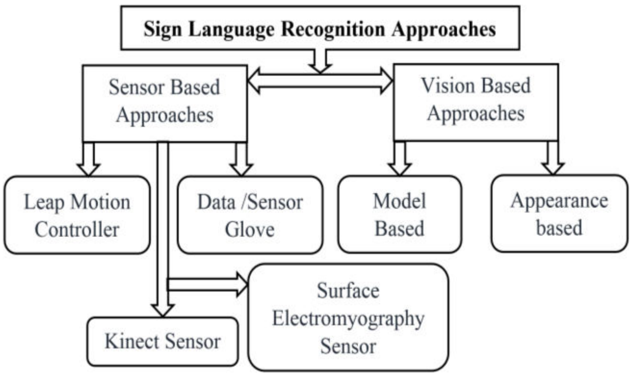

Fig 3. Sign Language Recognition Approaches

4.1 Sensor based approaches

These approaches uses sensors of various types and external hardware to collect the data of gesture done. Then the collected data is analyzed and decisions are represented according to recognition model. In Hand Gesture Recognition (HGR), many sensors are placed on the hand to make hand gesture, once hand gesture is done then the data is recorded for further analysis. Main drawback of this approach is that the complex gestures cannot be performed [3]. Further sensor based approaches are classified as follows:

4.1.1 Leap Motion Controller approach

It is a small USB peripheral sensor (Fig 4), detects the movement of hand and translates the detected signal into computer commands [4]. It contains three infrared (IR) LED's and two IR cameras. Camera generates 300 frames per second of reflected data and IR light generated by LED. These signals are send to the computer over USB cable for processing. Ching-Hua Chuan et al [5] present an ASLR system using leap Motion Sensor, provides robust and economical result comparing to cyber glove sensor and Kinect sensors.

4.1.2 Kinect Sensor approach

It is a motion sensor with Xbox 360 (Fig 4) gaming console [4] which has depth sensor (consist IR laser), RGB camera and Multi-array microphone. It support 3D motion capture of complete body. Also recognizes facial movement and voice too. Cao Dong et al describes a method for 24 ASL alphabet recognition, very useful in controlling industrial robots, factory floors, and hospital [6].

Fig 4. (a) Leap Motion Sensor, (b) Kinect for Xbx360

4.1.3 Data/Sensor Glove approach

In this approach, signer needs to wear a sensor glove (Fig 5) to detect finger and hand gesture signal. This analog signal is converted into digital form by ADC. Data glove also contains flex sensor and accelerometer [4]. Flex sensor detects the bend signal. Major drawbacks of this approach is that the signer needs to wear the sensor glove, which will lead less natural behavior and sensor gloves are pretty expensive.

Dhananjai Bajpai et al [7], focused on ASL recognition glove for images text and speech display on cell Phone with 2-way communication using only 3 contact sensors out of 5 contact sensors in the hand glove. This simplifies the complexity in hardware. Kiratey Patil et al, designed a sensor glove system for capturing ASL gestures, converts them to corresponding English letter and displayed on LCD [8]. A hand glove with flex sensor translator, to translate ASL alphabets and number, also some single hand ISL gestures is developed in [9].

4.1.4 Surface Electromyography Sensor approach

The sEMG sensor (Fig 5) will generate distinct muscle electrical patterns [10] using various signs/gestures performed by the signer. For SLR, if sEMG sensors are located on the forearm, then it classifies different finger movements and hand gestures clearly. This approach can be used to spot the neuromuscular diseases and to study human kinetics. Key factor EMG signals are that, they are not affected [11] by light illumination but having plenty of wires around the arm is a difficulty for acquiring sEMG signal. Jian Wu et al [10], proposed wearable Inertial Measurement Unit (IMU) and sEMG sensors to detect hand gestures in ASLR. Celal Savur et al [11], presented a real time ASLR for 26 English alphabets. In [12], a real time ASLR for 40 commonly used daily conversation words is proposed.

Fig 5. (a) Data Glove with Flex Sensor, (b) sEMG Sensor with cables

Various research works are carried out and also going on by several researchers on sensor based ASLR. Few of those research works are highlighted in the following table followed by a recognition rate comparison graph.

Table 1. Highlights of Sensor Based ASLR Literature Review

Sl #	Hardware used	Year of Public-ation	Type of Gest-ure	Results / Remarks
1	sEMG and IMU [10]	2016	Real Time HGR	Avg. accuracy **90.7%** for 80 commonly used ASL.
2	Microsoft Kinect and Color Glove [6]	2015	Static HGR	Accuracy **92%** for 24 static ASL alphabets. Used in industrial robots, factories, and hospital.
3	sEMG and Bio Radio 150 [11]	2015	Off-line and Real Time HGR	Avg. accuracy **86.7%** for offline and real time system. Drawback: Lot of wire connection prevents to perform a gesture easily.
4	sEMG and Wrist-Worn Motion [12]	2015	Real time HGR	Accuracy **95.94%** for 40 commonly used words of daily conversation and 4 subjects.
5	Hand Glove, AVR Microcont-roller [7]	2014	HGR	Accuracy **83%** for 36 gestures with 2 way communica-tion between 2 cell phones. Uses 3 flex sensors out of 5 flex sensors, which simplifies the complexity of hardware
6	Leap Motion Sensor [5]	2014	HGR	Avg. Accuracy **76.3%** for 26 ASL alphabets using k-NN and SVM methods

Fig 6 shows the recognition rate comparison graph of various ASLR techniques represented in Table 1

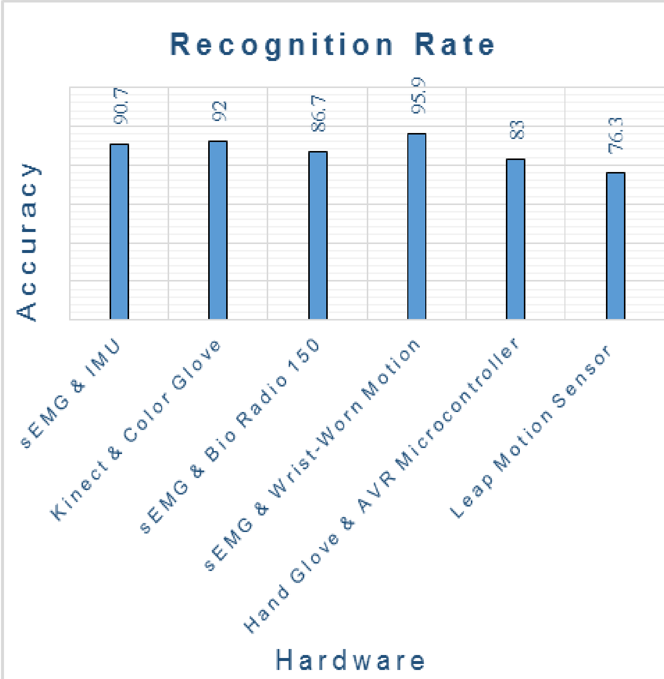

Fig 6. Recognition Rate of Various Sensor Based ASLR Techniques

4.2 Vision based approaches

This approach uses the web camera to capture the hand gestures and facial expressions from the signer. Image processing (IP) algorithms (consist various IP operations like background subtraction, segmentation, skin region detection, features extraction, classification and recognition) are need to be developed to recognize the captured gestures and facial expressions. This approach is quiet easier to the signer as it is not necessary to wear any hardware like sensor glove but related to IP algorithms there may be an accuracy drawback. Further vision based approaches are classified as follows [13]:

4.2.1 Model based approach

This approach uses 3D information of fundamental elements of the body parts. These information are very useful to obtain various key parameters like joint angles, palm position and so on. Skeletal and/or volumetric models are used in this approach as these models are better suited in computer vision and computer animation industry.

4.2.2 Appearance based approach

Images or videos are used as inputs in this approach as they don't use spatial illustration of the body. Using a template database, the parameters are derived from the images or videos directly. Few templates are depending on the deformable 2D templates of the human body parts, mainly hands. Sets of points on the outline of an object are referred as deformable templates, used as interpolation nodes for the outline approximation of an object. An approach implemented in [14] is very useful for learners to practice sign language.

5. TECHNIQUES OF VISION BASED SLR PROCESS

Following are the some of the significant techniques and related works of Vision based SLR process.

5.1 Feed Forward Artificial Neural Network (FF ANN)

In FF ANN, connections between the units do not form a cycle. As such, it is dissimilar from Recurrent Neural Networks [16]. The FF ANN was the foremost and simplest type of ANN invented. Here, the information moves in single direction, forward from the input nodes to the output nodes through the hidden nodes. 37 ASL signs and numbers are recognized with 30 feature vectors using FF ANN and back propagation NN algorithms [17].

5.2 Backpropagation NN (BPNN)

BPNN is a common method of training ANNs and used in conjunction with the gradient descent optimization method. The algorithm recaps a two phase cycle i.e., propagation and weight update. Suppose an input vector is provided to the network, it is propagated forward, layer by layer, till it reaches the output layer. Then the network output is compared to the output desired, using a loss function, and calculates error value for each of the neurons [18]. The error values are propagated backwards from the output, till each neuron has an associated error value that approximately represents its contribution to the original output. In [17], backpropagation algorithm is used for training the Neural Network.

5.3 Speeded Up Robust Features (SURF)

SURF is a patented detector for recognizing local feature and descriptor used for tasks i.e., people or faces, image registration, classification or 3D reconstruction, object recognition, to objects tracking and to extract interest points [19]. SURF uses the determinant integer approximation of Hessian blob detector to detect interest points, which can be calculated with three integer operations by a recalculated integral image. Feature descriptor is depends on the response of the sum of the Haar wavelet around the interest point, which can also be calculated with the aid of the integral image. Herbert Bay et al [19], present a novel detector-descriptor method, which provides distinctiveness, much faster and robust recognition.

5.4 Convexity Defect (CD)

Convexity Defects is a feature extraction technique, used to gesture recognition by calculating silhouettes of subjects using a morphological algorithm [20]. It is a cavity in an object segmented out from an image. This means a region that do not belong to the object but placed inside of its outer boundary. In 2015, a real time HGR process was implemented using codebook algorithm for background subtraction, later used contour, convex haul and convexity defect calculation [21].

5.5 Convex Hull

Convex envelope (hull) of set X of points in a Euclidean plane or in a Euclidean space is the tiniest convex set, which contains X [22]. When X is a bounded plane subset, the convex hull may be picturized as the shape enclosed by a rubber band stretched around X. Formally, the convex envelope may be defined as the intersection of all convex sets containing X. Convex hulls may be prolonged from Euclidean spaces [23] to arbitrary real vector spaces. Further they may be generalized. Concept is utilized in [21].

5.6 K-Curvature (K-C)

Curvature is a quantity by which a geometric object such as a surface diverges from existence a flat plane, or a curve from being straight as in the case of a line, but this is well-defined in various ways depending on the context. K- Curvature is any of a number of loosely connected conceptions in various parts of geometry. There is a major dissimilarity between extrinsic curvature that is defined for embedded objects in an another space in a way which relates to the circles' curvature radius, which touch the object and intrinsic curvature, which is defined in terms of the lengths of curves within a Riemannian manifold. Usually Curvature is a scalar quantity [24], but one may also state a vector of curvature, which takes into an account the way of the bend in addition to its magnitude. In [17], a feature extraction algorithm is implemented using k-curvature algorithm by combing with convex hull algorithm.

5.7 HSV Color Model

It is a perceptual color model, which describes color based on 3 elementary features of the color i.e., Hue (H) is a basic feature of color such as red, yellow, green ranging from 0 to 360 in a color space. Saturation (S) defines the purity of the color. In a color space, S specifies the amount of grey color, ranging up to 0 to 100% or 0 to 1, where 0 is grey and 1 is primary color. Value (V) is the brightness (Luminance) of the color which changes saturation ranging between 0 to 100% where 0 is totally black and 100 is pure white in color space. This model permits us to state the skin pixel boundary only in terms of hue and saturation. A. Sharmila Konwar et al [25], developed an ASLR system using canny edge detection technique and HSV color model with the help of set of morphological operations.

5.8 Morphological Operation

Morphology in mathematics is a tool for mining the useful image components in representation and description of region shape, i.e., skeletons, boundaries, and convex hulls. The language of mathematical morphology is set theory, and used apply directly to binary images. A point is either in the set or reset, and the usual operators of set can be applied to them. The two basic operations on sets: an image, and the structuring element. In [25], an erosion and dilation is used to get a thin structure of the edge detected binary image for ASL gesture recognition.

5.9 Hidden Markov Model (HMM)

HMM is a Statistical Markov Model (SMM), system being modeled is assumed to be a Markov process with hidden (unobserved) states. In HMM, the state is not directly visible, but the output is visible and output is state dependent. Over the possible output tokens, each state has a probability distribution. So, an HMM generated sequence of tokens provides some information about the sequence of states [26]. HMMs are very useful in speech, gesture recognition, musical score following, handwriting, partial discharges and bioinformatics. LumKin Yun et al [27], presents a system for recognizing ASL character using HMM as classifier with better classification rate.

5.10 Principal Component Analysis (PCA)

PCA is a statistical technique, uses an orthogonal transformation to translate the set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal. The number of principal components is less than or equal to the lesser of the number of original variables or number of observations [28]. So that the first principal component has the biggest possible variance, and each subsequent component in turn has the utmost variance possible under the limitation that it is orthogonal to the earlier components. The resultant vectors are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables and used as a tool in exploratory data analysis and for making predictive models as well. PCA technique is used in [25] for feature extraction of gestures in ASLR.

5.11 Kalman Filtering (KF)

Kalman filtering is a recursive algorithm, run in real time, using only current input measurements and earlier calculated state and its matrix of uncertainty. KF uses a series of measurements observed over time, holding statistical noise and other inaccuracies, and produces estimates of unidentified variables that tend to be more precise than those based on a single measurement alone, by using Bayesian inference and approximating a joint probability distribution over the variables for each timeframe. Here, among many applications, navigation, guidance, signal processing, econometrics, and control of vehicles, particularly spacecraft [29] and aircraft are some common applications. Herve Lahamy et al [30], presented a rotation invariant 3D hand gesture of ASL alphabets using Kalman Filter for tracking of hand moments.

5.12 Fuzzy Decision Trees (FDTs)

The fuzzy rules, were generated were 'flat' (applied concurrently). Fuzzy rules creates more sense than standard bivalent rules, the usual extension is to yield a clue from decision trees and an effort to build FDTs [31]. Rather than bivalent, these would be identical to standard decision trees excluding that the decisions at each branching point would be fuzzy. This makes computing the finest split somewhat rigid, but would be much rewarding by allowing smaller trees with fewer leaves and internal nodes to encapsulate a richer amount of information. Gaolin Fang et al [32], proposed a FDTs and heterogeneous classifiers for huge set of vocabulary of SLR, where the recognition time is 11 time faster and accuracy is 0.95% more over HMM.

5.13 Discrete Cosine Transform (DCT)

DCT is strongly associated to the Discrete Fourier Transform (DFT) uses only real numbers so the result is real. DCT inclines to focus information, making it beneficial for several science and engineering for lossy image and audio compression applications and serving in minimalizing feature vector size for dissimilar applications as well. DCTs are approximately equivalent to DFTs of twice the length, functioning on real data with even symmetry, but in some variations the input and/or output data are moved by half a sample. In [33], a simple and effective system is proposed for hand gesture recognition in non-complex background using DCT, Haar, Walsh, and Kekres transforms. DCT provides greater recognition rate than PCA.

5.14 K-Nearest Neighbors (K-NN)

K-NN is the simplest and easiest to implement machine learning algorithm, used for classification and regression. In both the cases, this technique is based on the nearby training samples in the feature space. Distance among the test sample and training samples are calculated first using Euclidean distance when the test sample is given. Later, least distance 'k' nearest neighbours are determined. The test sample is categorized based on the mainstream votes of the nearest neighbours whenever the nearest neighbours are found. Keeping the input feature vectors and their equivalent labels are involved by k-NN Training. A non-labeled query image is simply allotted to the label of its k nearest neighbours in testing phase. S. Nagarajan et al [34], proposed a 10 ASL numeric hand gestures (0-9) recognition using k-NN with comparable classification accuracy of 96% for k=5.

5.15 Support Vector Machines (SVMs)
SVM is a powerful supervised learning method, for analyze the data in classification and regression process. Also used to categorize the test data group as one of the many signs, based on the value of the feature. Each sample in a training samples set is marked as either belonging to classification and regression. SVMs provision vectors and kernels are working for several learning tasks. Various tasks can perform in different areas by selecting suitable kernel methods. SVM creates a separating hyper plane in a high dimensional space. An ASLR classification and recognition for 10 numeric (0-9) hand gestures [34] is proposed using SVMs and obtained better accuracy than k-NN.

5.16 Probabilistic Neural Network (PNN)
PNN is a feedforward NN and radial-basis function network, widely used in classification and recognition. PNN is derived by Bayesian network, is contains an input layer, radial basis layer and competitive layer. The competitive layer classifies depending on radial basis unit with leading output. A radial basis function (kernel function) is calculated by first layer, for distances between the input and each of the kernel vectors. The radial distance is an argument for the function so the name of the function is radial basis function. The weight is calculated as its radial distance function. The K sums of the radial basis kernel outputs and biggest sum is find by the second layer. The biggest sum class has the extreme probability of being true and is taken as network output. PNN has less training time. Vinitha K V et al [35] present a system for detection of face using PNN architecture with eclipse fitting algorithm.

5.17 Cartesian Genetic Programming (CGP)
CGP is a very simple and well-known form of Genetic Programming (GP). CGP is depending on the feed-forward digital circuits' evolutionary design. Here, genes are signified by nodes, which has some features and aspects [36]. In CGP system, a node has a function that implements principally few inputs and an output and an array of nodes in middle. The traditional form of CGP has undergone a number of improvements that made it more beneficial, effectual and flexible in several ways, which include cyclic connections (recurrent-CGP), Self-Modifying CGP (SMCGP), encoding ANNs and automatically defined functions [37]. SMCGP uses functions that cause the grown programs to change themselves as a time function. This makes it possible to find generalized solutions to mathematical algorithms and classes of problems. Recurrent-CGP permits evolution to build programs that consists cyclic and acyclic, connections. This allows application to tasks that need inner states or memory, and create recursive equations. In [36], presented an ASLR system for 26 English alphabets and provides faster recognition rate than conventional Genetic programming.

5.18 Self-Organizing Map (SOM)
SOM, another type of ANN used for dimensionality reduction i.e., trained using unsupervised learning to yield a low-dimensional (typically 2-D), discretized illustration of input space of the training samples, called a Map. In SOM, instead of mining the features from image for training and recognition, an image itself is transf0rmed to radial walled contour and trained. SOM is assured for user free recognition of gesture. Like most ANNs, there are two modes in SOM: 'Training' builds the map using input data, whereas 'mapping' automatically classifies a new input vector. Nagaraj N. Bhat et al [38], presented a SOM based ASLR system for 18 hand gesture of ten subjects and the method working fine in uncontrolled conditions with effective time saving.

5.19 Edge Orientation Histogram (EOH)
EOH could be computed capably by integral images and, are robust for minor variants in rotation and position. It highlights the use of local indicators constitution in the same manner as weak classifiers for AdaBoost allied detection techniques [39]. EOH matches with local object regions. In [40], an ASL recognizer is presented for real-time static alphabet with frame size 160x120 using web cam of distance 1 meter attached to laptop.

Various research works are carried out and also going on by several researchers on vision based ASLR. Few of those research works are highlighted in the following table followed by a recognition rate comparison graph.

Table 2. Highlights of Vision Based ASLR Review

Sl #	**Techn-ique used**	**Year of Publication**	**Type of Gest-ure**	**Results / Remarks**
1	FF ANN [17]	2017	Real Time HGR	Accuracy **94.32%** for 37 signs of ASL alphabets and numbers. 1850 sample images captured using mobile video camera in black background. Future Scope: Hand movement detection for word recognition.
2	EOH [40]	2016	Real Time static HGR	Accuracy **88.2%** with time 0.5 secs for 100 ASL symbols out of 2600 training samples in Complex backgrou-

				nd with Natural lighting. Signs captured using 10 mega pixel E-cam fixed to laptop with 1 meter distance.
3	SURF , and SRG [41]	2016	Real Time HGR	Average accuracy **97.13%** for 16 various ASL gestures. SL may be converted to text and speech outputs on the smartphone app. Future Scope: Enlarge to a wider vocabulary. Improve the processing speed and accuracy.
4	Ada, and Haar [42]	2016	Real Time HGR	Accuracy **98.7%** for 28000 samples. Input: Live video in complex background. Output: Text and speech. VGA 640x480 resolution.
5	SVM [34]	2015	Static HGR	Accuracy **98.4%** for ASL numbers which outperforms the k-NN for the same dataset. Future Scope: Dynamic gesture from real time video.
6	k-NN [34]	2015	Static HGR	Accuracy **96.6%** for ASL numbers
7	BPNN [34]	2015	Static HGR	Accuracy **93.01%** for ASL numbers
8	CD [43]	2014	Real Time HGR	Accuracy **80%.** Sensor distance from 80-175 cm.
9	K-C [43]	2014	Real Time HGR	Accuracy **92%.** Sensor distance from 80-175 cm.
10	HMM [27]	2013	Real Time HGR	Accuracy **93.75%** for ASL alphabets. Devel -oped using Visual Studio 2010. Gesture capturing time 2 seconds.
11	DCT [33]	2013	HGR	Accuracy **100%** for 48 (24*2) images are tested. Taken 120 (24*5) images in **plane background**

				with different angle and position for training. Transform of sizes 4x4, 8x8, 16x16, 32x32, and 64x64 are used. Future Scope: complex gestures in complex background
12	Hu moment [14]	2013	HGR	Accuracy **76%** without a controlled background with small light adjustments. Useful for learners to practice sign language. Sign to text and speech
13	SOM [38]	2013	HGR	Average accuracy of **92%** of 18 ASL gestures of 10 subjects. Future scope: More efficient and robust ASLR in real time environments using stereo vision.
14	KF [30]	2012	Real Time HGR	Accuracy **93.88%** for 14732 samples of 12 postures of ASL 3D alphabets. Future Scope: Improvement of noise removal, improving the alignment of 2-hand postures.
15	CGP [36]	2011	Real Time HGR	Average Accuracy **90%** for 26 signs of ASL alphabets. Future Scope: Improve the accuracy and speed.
16	NN [44]	2010	Real Time HGR	Accuracy **89%** 320x240 resolution size hand gestures in uniform background and normal illumination. Future Scope: Detect moving hand and recognize complex gestures
17	FDT [32]	2004	Real Time HGR	Accuracy **83.7%** of 5113 sign vocabulary with an average recognition time 0.263 sec/word. Future

				Scope: Recognition of Gaze, facial Expressions, Mouth Movements, position, and motion of the trunk and head.

Fig 7 shows the recognition rate comparison graph of various SLR techniques represented in Table 2

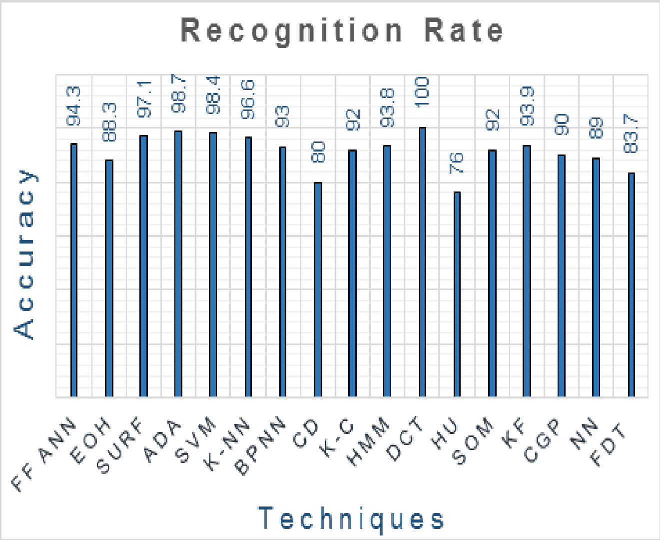

Fig 7. Recognition Rate of Various Vision Based ASLR Techniques

Note: In Fig 7, SURF = Speeded Up Robust Features and Seeded Region Growing, ADA = Adaboost and Haar Classifiers, HU=Hu Moment.

Recognition Rate shown above is purely depends on different background, distance from the signer to camera, number of sample trained, and number of samples tested.

Fig 8 depicts recognition rate of Top 3 Sensor based and Vision Based ASLR Techniques.

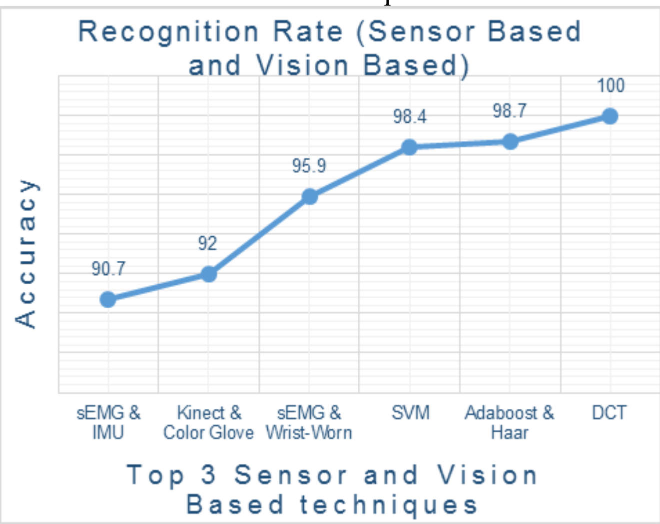

Fig 8. Recognition Rate of Top 3 Sensor Based and Vision Based ASLR Techniques

Fig 9 spectacles the average recognition rate of Sensor Based and Vision Based ASLR techniques

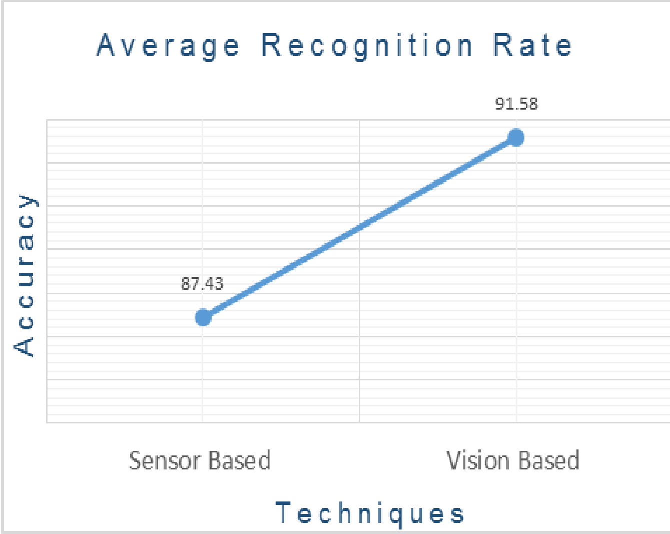

Fig 9. Average Recognition Rate of Sensor Based and Vision Based ASLR Techniques

Fig 10 sights the average recognition rate of various techniques in Sensor Based and Vision Based ASLR system.

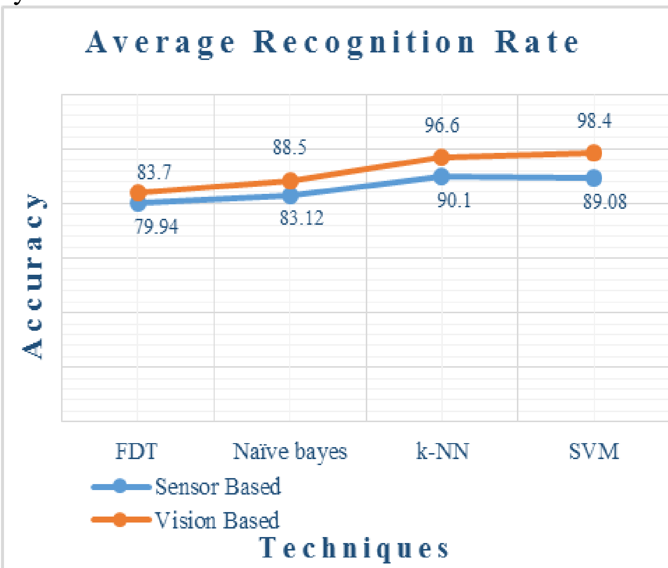

Fig 10. Average Recognition Rate of various techniques in Sensor Based and Vision Based ASLR system.

6. CONCLUSION

This review paper presents the ASLR process, and various constraints of ASLR system. Also in this paper, a sincere effort has been put to present various sensor and vision based ASLR techniques like sEMG, Hand Glove with flex sensors, Kinect, Leap Motion Sensor, CGP, SOM, PCA, FF ANN, FDT, SVM, Naïve Bayes, k-NN, and many. From the above study it is clear that the vision based gesture recognition has made outstanding progress in the field of gesture recognition of ASLR. One of the benefit of vision based method is that it can be implemented practically and no need to wear or use any sensors or hand gloves. In order to enhance the accuracy of the gesture recognition and hence SLR, need more work to be done in the field of feature extraction and classification. In vision based ASLR techniques, DCT provides 100% recognition rate but it is trained and tested on very limited samples in plane background. Next to DCT, the combination of Adaboost and Haar classifiers offers recognition rate of 98.7% which is tiny better than the SVM technique which provides 98.4%. However, many of the vision based ASLR techniques offers over 90% of recognition rate.

REFERENCES

[1] https://www.nidcd.nih.gov/health/American-sign-language

[2] https://incubator.wikimedia.org/wiki/Category: Incubator:Test_wikis_of_sign_languages

[3] Vaibhavi S. Gandhi, Akshay A. Khond, Sanket N. Raut, Vaishali A. Thakur, Shabnam S. Shaikh, A Review of Various Gesture Recognition Techniques, *International Journal of Engineering And computer Science*, 2014, 8202-8206

[4] Manisha U. Kakde, Mahender G. Nakrani, Amit M. Rawate, A Review Paper on Sign Language Recognition System For Deaf and Dumb People using Image Processing, *International Journal of Engineering Research & Technology*, 2016, 590-592

[5] Ching-Hua Chuan, Eric Regina, Caroline Guardino, "*American Sign Language Recognition Using Leap Motion Sensors*", 13th International Conference on Machine Learning and Applications, 2014, pp.541-544

[6] Cao Dong, Ming C. Leu and Zhaozheng Yin, "*American Sign Language Alphabet Recognition Using Microsoft Kinect*", 2015 IEEE, pp.44-52

[7] Dhananjai Bajpai, Uddaish Porov, Gaurav Srivastav, Nitin Sachan, "*Two Way Wireless Data Communication and American Sign Language Translator Glove for Images Text and Speech Display on Mobile Phone*", 5th International Conference on Communication Systems and Network Technologies, 2015, pp.578-585

[8] Kiratey Patil, Gayatri Pendharkar, Prof. G. N. Gaikwad, American Sign Language Detection, *International Journal of Scientific and Research Publications*, Volume 4, Issue 11, 2014, 1-6

[9] Sunita V. Matiwade, DR.M.R.Dixit, Electronic Support System to Interpret Sign Language of Communication used by Deaf and Dumb Community, *Journal of Emerging Technologies and Innovative Research*, 2015, 52-57

[10] Jian Wu, Lu Sun, Roozbeh Jafari, A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors, *IEEE Journal of Biomedical And Health Informatics*, 2016, 1281-1290

[11] Celal Savur, Ferat Sahin, "*Real-Time American Sign Language Recognition System by Using Surface EMG Signal*", 2015 IEEE, pp.497-502

[12] Jian Wu, Lu Sun, Leonardo Estevez Roozbeh Jafari, "*Real-time American Sign Language Recognition Using Wrist-worn Motion and Surface EMG Sensors*", 2015 IEEE

[13] Suriya M, Sathyapriya N,Srinithi M,Yesodha V, "*Survey on Real Time Sign Language Recognition System: An LDA Approach*", International Conference on Explorations and Innovations in Engineering & Technology, 2016, pp.219-225

[14] Matheesha Fernando, Janaka Wijayanayaka, "*Low cost approach for Real Time Sign Language Recognition*", 2013 IEEE, pp.637-642

[15] Archana S. Ghotkar, Dr.Gajanan K. Kharate, Study of Vision Based Hand Gesture Recognition Using Indian Sign Language, *International journal on smart sensing and intelligent systems*, 2014, 96-115

[16] https://en.wikipedia.org/wiki/Feedforward_ neural_network

[17] Md. Mohiminul Islam, Sarah Siddiqua,, Jawata Afnan, "*Real Time Hand Gesture Recognition Using Different Algorithms Based on American Sign Language*", 2017 IEEE

[18] https://en.wikipedia.org/wiki/Backpropagation

[19] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, Speeded Up Robust Features, *ETH Zurich, Katholieke University*, Leuven

[20] https://stackoverflow.com/tags/convexity-defects/info

[21] Manasa Srinivasa H S, Suresha H S, Implementation of Real Time Hand Gesture Recognition, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol 3, Issue 5, 2015, 4060-4065

[22] Andrew, A. M, Another efficient algorithm for convex hulls in two dimensions, *Information Processing Letters*, 1979, 216–219

[23] Brown, K. Q, Voronoi diagrams from convex hulls, *Information Processing Letters*, 1979, 223–228

[24] https://en.wikipedia.org/wiki/Curvature

[25] A. Sharmila Konwar, B. Sagarika Borah, C. Dr.T.Tuithung, "*An American Sign Language Detection System using HSV Color Model and Edge Detection*", International Conference on

Communication and Signal Processing, 2014, pp.743-747

[26] Thad Starner, Alex Pentland, "Real-Time American Sign Language Visual Recognition from Video Using Hidden Markov Models. Master's Thesis", *MIT*, 1995

[27] Lum Kin Yun, Tan Tian Swee, Rina Anuar, Zuraimi Yahya, Azli Yahya, "*Sign Language Recognition System using sEMG and Hidden Markov Model*", Recent Advances in Mathematical Methods, Intelligent Systems and Materials, 2013, pp.50-53

[28] Rafael C Gonzalez, and Richard E Woods. "*Digital Image Processing*", 2nd Edition,

[29] Paul Zarchan, Howard Musoff, *"Fundamentals of Kalman Filtering: A Practical Approach*", American Institute of Aeronautics and Astronautics, Incorporated, 2000

[30] Herve Lahamy, Derek D. Lichti, "*Towards Real-Time and Rotation-Invariant American Sign Language Alphabet Recognition Using a Range Camera*", Department of Geomatics Engineering, The University of Calgary, Canada, 2012, pp.14416-14441

[31] http://www.kev.pulo.com.au/ai/fuzzyml_report/node17.html

[32] Gaolin Fang, Wen Gao, Debin Zhao, "*Large Vocabulary Sign Language Recognition Based on Fuzzy Decision Trees*", 2004 IEEE, Vol. 34, No. 3, pp.305-314

[33] Dr. Tanuja K. Sarode, Vaishali Sakpal, Performance Comparison of Transform Techniques for Hand Gesture Recognition, *International Journal of Advanced Research in Computer Science*, Vol 4, No. 9, 2013, 245-250

[34] S. Nagarajan, T. S. Subashini, Image Based Hand Gesture Recognition using Statistical Features and Soft Computing Techniques, *International Journal of Emerging Technology and Advanced Engineering*, Vol 5, Issue 7, 2015, 476-482

[35] Vinitha K V and G Santhosh Kumar, "*Face recognition using probabilistic Neural Network*", World Congress on Nature & Biologically Inspired Computing, 2009, pp.1388-1393

[36] Fahad Ullah, "*American Sign Language Recognition System for Hearing Impaired People Using Cartesian Genetic Programming*", 5th International Conference on Automation, Robotics and Applications, 2011, pp.96-99

[37] Julian Miller, Andrew Turner, *GECCO Companion '15*, Proceedings of the Companion Publication of the Annual Conference on Genetic and Evolutionary Computation, 2015, pp.179-198

[38] Nagaraj N. Bhat, Y V Venkatesh, Ujjwal Karn and Dhruv Vig, "*Hand Gesture Recognition using Self Organizing Map for Human Computer Interaction*", International Conference on Advances in Computing, Communications and Informatics, pp.734-738, 2013

[39] Bram Alefs, Guy Eschemann, Herbert Ramoser, Csaba Beleznai, "*Road Sign Detection from Edge Orientation Histograms*", 2007 IEEE, pp.993-998

[40] Jayshree R. Pansare, Maya Ingle, "*Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram*", International Conference on Image, Vision and Computing, 2016, pp. 86-90

[41] Cheok Ming Jin, Zaid Omar, Mohamed Hisham Jaward, "*A Mobile Application of American Sign Language Translation via Image Processing Algorithms*", 2016 IEEE, pp.104-109

[42] Vi N.T. Truong, Chuan-Kai Yang, Quoc-Viet Tran, "*A Translator for American Sign Language to Text and Speech*", 2016 IEEE

[43] Marek Vaneo, Ivan Minarik, Gregor Rozinaj, "*Evaluation of Static Hand Gesture Algorithms*", 21st International Conference on Systems, Signals and Image Processing, 2014, pp.83-86

[44] G.R.S. Murthy, R.S. Jadon, "*Hand Gesture Recognition using Neural Networks*", 2010 IEEE, pp.134-138