International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT217255

Accessible System for Sign Language Computation

Patel Harshil[1], Patel Jay[1], Patel Smit[1], Thakkar Mayur[1], Dr. Sheshang Degadwala[2]

[1]U.G, Department of Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India
[3]Associate Professor, Department of Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India

Article Info
Volume 7, Issue 2
Page Number: 222-230

Publication Issue :
March-April-2021

Article History
Accepted : 05 April 2021
Published : 11 April 2021

ABSTRACT

Sign Language is most accepted and purposeful way of communication for deaf and mute folks of the society. signing uses gestures, head/body movements and facial expressions for communication. it's a strong manner of Non-Verbal communication among humans. each country has its own developed signing. The language which is employed in Bharat is named as "Indian Sign Language (ISL)". terribly lessresearch work has been carried out during this space as ISL has standardized recently. presently several researchers have primarily focused on gesture recognition that has been
recorded underneath static hand gesture. solely few works have been reported for recognition of dynamic hand gesture. several ways are developed to recognize alphabets and numbers of ISL. The major steps concerned in planning such a system area unit gesture acquisition, trailing and segmentation, feature extraction and gesture recognition. This paper presents a survey on varied hand gesture recognition approaches for ISL.

General Terms : Hand Gesture Recognition, Human computer interaction, YOLO3

Keywords : Sign language recognition, Indian Sign Language, ANN, CNN, OpenCV, YOLO3

I. INTRODUCTION

Sign language provides the way for the deaf and mute people to speak with the skin world. Instead of voice, language uses gestures to communicate. language is associate organized method of communication during which each word or alphabet is assigned to a selected gesture and that means [2].

Sign language is created from a spread of gestures produced by varied facial expressions and movements of hands or head/body. within the last several years there has been associate exaggerated interest among the researchers within the field of language recognition to introduce a way of interaction from human – human to human – laptop interaction. varied applications of Gesture Recognition System are: Human laptop Interface, Video vice, increased reality, Home appliances, Robotics, language, etc. Hearing impaired and speech impaired folks rely upon sign language interpreters for communication. But finding masterly and trained interpreters for his or her day to day interactions throughout life time could be a terribly

difficult job and additionally too valuable [7]. everywhere the world, a range of sign languages exist. The sign language depends on the traditions and spoken language of that place. Indian language (ISL) is used by hearing and speech impaired folks in India [3]. The gestures square measure primarily divided into two classes: Static gestures and Dynamic Gestures. Static gestures embrace solely configurations and poses whereas dynamic gestures embrace postures, strokes, pre-strokes, phases and additionally emotions [1]. ISL

alphabets and numeric signs square measure portrayed in Fig.1 and Fig.2 severally

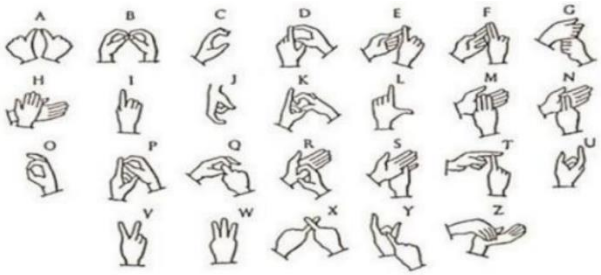

Fig.-1: Representation of ISL Alphabets (A-Z) [3]

II. SIGN LANGUAGE RECOGNITION-DIFFERENT APPROACHES

Sign language recognition is a vital application of gesture recognition. There square measure principally 2 totally different approaches in signing recognition - Glove primarily based} approach and vision based approach. the primary class needs signers to wear a detector glove or a coloured glove. The carrying of the glove simplifies the task of segmentation throughout process. the disadvantage of this approach is that the signer should wear the detector hardware beside the glove throughout the operation of the system. Vision primarily based approach uses image process algorithms to discover and track hand signs also as facial expressions of the signer. This approach is less complicated to the signer since there's no have to be compelled to wear any additional hardware. However, there square measure accuracy issues associated with image process algorithms and these issues square measure nevertheless to be changed. There square measure once more 2 totally different approaches in vision {based|based mostly|primarily primarily based}} signing recognition: 3D model based and look based [1] 3D model primarily based strategies build use of 3D data of key components of the body elements. victimisation this data, many vital parameters, like palm position, joint angles etc., is obtained. This approach uses volumetrical or skeletal models, or a mix of the 2. In pc animation trade and for pc vision functions, volumetrical approach is healthier suited. This approach is incredibly procedure intensive and additionally, systems for live analysis square measure still to be developed. Appearance-based systems use pictures or videos as inputs. They directly interpret from these videos/images. They don't use a spatial illustration of the body. The parameters square measure derived directly from the photographs or videos employing a templet info. Some templates square measure the deformable second templates of the human elements of the body, significantly hands. Deformable templates square measure sets of points on the define of associate degree object, used as interpolation nodes for the object's define approximation. one among the best interpolation functions is linear. It performs a mean form from purpose sets, purpose variability parameters and external deformations. These template-based models square measure largely used for hand-tracking, however might even be used for straightforward gesture classification. A second approach in signing gesture detection victimisation appearance-based models uses image sequences because the gesture templates. Either the photographs themselves, or sure options derived from these pictures is used because the parameters.

III. OVERVIEW OF SIGN LANGUAGE RECOGNITION SYSTEM

A simple diagram of a symptom language recognition system is shown in Figure one. the complete recognition method may be divided into 2 phases- coaching and testing. within the coaching part, the classifier has got to be trained victimization the coaching dataset. The info may be either created by the investigator himself or associate on the market info may be used. associate external digital camera, camera or integral digital camera within the laptops may be accustomed capture the coaching pictures. Most of the language recognition systems classify signs per hand gestures solely or in alternative words, facial expressions ar excluded. The necessary steps concerned in coaching part ar creation of info, preprocessing, feature extraction and coaching the classifier. The testing part includes video/image acquisition (input may be videos or images), preprocessing, feature extraction and classification.

3.1 Preprocessing

A preprocessing step is applied on the coaching pictures to extract the region of interest (ROI). The ROI are often hands if solely hand gestures square measure thought-about or each face and hands if the facial gestures are enclosed. typically the preprocessing step consists of filtering, image sweetening, image resizing, segmentation and morphological filtering. Filtering and image sweetening are often anybody of the ordinarily used ways. For segmentation, the algorithmic program that higher suits the input video/images needs to be selected . Otsu's thresholding [2], Background subtraction [3], complexion primarily based} segmentation [4] and motion based segmentation [5] are the ordinarily used segmentation techniques. throughout testing part, the take a look at pictures or videos are preprocessed to extract the region of interest.

3.2 Feature Extraction

Feature extraction is one in every of the foremost crucial steps of linguistic communication recognition since the inputs to the classifier ar the feature vectors obtained from this step. The techniques used for feature extraction ought to realize shapes dependably and robustly regardless of changes in illumination levels, position, orientation and size of the item in a very video/image. Objects in a picture ar delineated as assortment of pixels. For beholding we want to explain the properties of those teams of pixels. The options is obtained in numerous ways: wave decomposition [6], Haar wavelets, Haar-like options [7], texture options [8], orientation histogram[9], scale invariant feature transform[10], Fourier descriptors etc. In some cases the ROI pixels are used because the feature vector when a spatiality reduction victimization Principal element Analysis (PCA) [11]. The feature vector so obtained victimization anybody of the feature extraction strategies is employed for coaching the classifier.

3.3 Classification

A classifier is required in language recognition to classify the input signs into completely different categories. The feature vector obtained from the coaching info is employed to coach the classifier throughout the coaching part. once a take a look at input is given, the trained classifier identifies the category equivalent to the sign and displays the text or plays the sound. The take a look at inputs will be pictures or videos. most typically used classifiers area unit Hidden Andre Markoff Models (HMM), Artificial Neural Networks (ANN), Multiclass Support Vector Machines (SVM), Fuzzy systems, K

Nearest Neighbor (KNN) etc. The performance of the classifier is measured in terms of recognition rate.

3.3.1 Hidden Markov Models

A Hidden Andre Markoff model [20] may be a assortment of finite states connected by transitions. every state is characterised by 2 sets of probabilities: a transition chance and either a separate output chance distribution or continuous output chance density operate that, given the state, defines the condition chance of emitting every output image from a finite alphabet or a continuous random vector. The HMM approach to gesture recognition is intended by the booming application of Hidden Andre Markoff modeling techniques to speech recognition issues. HMM is a doubly stochastic model and is appropriate for handling the stochastic properties in gesture recognition. Instead of using geometric features, gestures square measure regenerate into serial symbols. HMMs are employed to represent the gestures, and their parameters square measure learned from the coaching knowledge. Based on the most likely performance criterion, the gestures can be recognized by evaluating the trained HMMs.

3.3.2 Artificial Neural Networks

An artificial neural network [16] involves a network of simple processing elements (artificial neurons) which can exhibit complex global behavior, determined by the connections between the process parts and part parameters. It consists of AN interconnected cluster of artificial neurons and processes info employing a connectionist approach to computation. In most cases AN ANN is AN reconciling system that changes its structure supported external or internal info that flows through the network throughout the educational part. The utility of artificial neural network models lies within the undeniable fact that they'll be accustomed infer a function from observations. There are many neural networking algorithms that will be used for gesture recognition. The different networks square measure feed forward networks, Elman neural networks, Self-organizing networks etc. There square measure many backpropagation algorithms offered for coaching the neural networks.

3.3.3 Support Vector Machine

The SVM [15] could be a in style pattern recognition technique with supervised learning. Since it divides the feature area for every category, the SVM will handle unknown knowledge well, though it's not suited to grouping sample knowledge. it's originally developed by Vapnik and colleagues at bell laboratories. it had been truly developed for determination binary call issues. the fundamental SVM takes a group of input file and predicts, for every given input, that of 2 potential categories forms the output. Thus, it is referred to as as a non-probabilistic binary linear classifier. For multi-class issues, such issues square measure rotten into many two-class issues that may be self-addressed directly mistreatment many SVMs. additionally to activity linear classification, SVMs will expeditiously perform non-linear classification mistreatment what's referred to as the kernel trick, implicitly mapping their inputs into high-dimensional feature areas. the various approaches for determination multi-class issues with SVM square measure one against all, one against one, call directed acyclic graphic approach etc

IV. ISL RECOGNITION

Figure two shows the signs in ISL resembling the English alphabets A, B, C, D and E. ISL recognition is relatively new within the field of language recognition. A sensible framework for ISL gesture primarily based human golem interaction had been planned in [12]. ISL video had been accustomed

compose many hand gestures to coach the HOAP-2 golem in real time. Orientation bar graph feature was extracted for real time classification victimisation geometrician distance metric. ISL video had been captured by choosing many dynamic gestures (i.e. sequence of frames) in real time victimisation digital camera. The classifier was trained victimisation twenty ISL gestures. The work was dole out in WEBOTS platform victimisation real time JAVA primarily based code (developed by them). The classified ISL gestures with average accuracy of ninetieth were entirely mapped with specific gesture primarily based applications on automaton golem. In [15], the authors took their dataset for dual-hand gesture from the twenty six alphabets of ISL. The dataset consisted a complete of 2340 (30×26× 3) pictures, three signers were requested to sign all 26 alphabets 30 times. Features were extracted from the images using Histogram of Orientation Gradient (HOG) and Histogram of Edge Frequency (HOEF). Finally, gestures were classified using Support Vector Machine (SVM). They had proposed a novel HOEF feature, using which they achieved a recognition rate of 98.1% of dual-handed gesture, and they have claimed in their paper that HOEF is better than HOG which is widely used by many for recognition.

A system had been proposed [14] by P.V.V. Kishore and P. Rajesh Kumar to automatically recognize gestures of ISL from a video stream of the signer. Their system converts words and sentences of ISL into voice and text in English. To accomplish the task, they had used powerful image processing techniques such as frame differencing based tracking, edge detection, wavelet transform and image fusion techniques to segment shapes in the videos. Then Elliptical Fourier descriptors were used for shape feature extraction and PCA for feature set optimization and reduction. Database of extracted features were compared with input video of the signer using a trained fuzzy inference system. The proposed system converts gestures into text and voice message with 91% accuracy. Around 80 gestures from 10 different signers were used. The entire system was developed in a user friendly environment by creating a Graphical User Interface (GUI) in MATLAB.

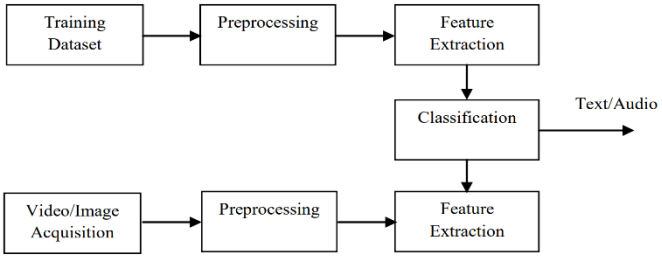

Figure 1. Simple Block Diagram of Sign Language Recognition System

P.V.V.Kishore and P.Rajesh Kumar [16] again proposed a real time approach to recognize gestures of ISL. The input video to the sign language recognition system was made independent of the environment in which signer was present. Active contours were used to segment and track the non-rigid hands and head of the signer. The energy minimization of active contours was accomplished by using object color, texture, boundary edge map and prior shape information. A feature matrix was designed from segmented and tracked hand and head portions. This feature matrix dimensions were minimized by temporal pooling creating a row vector for each gesture video. Pattern classification of gestures was achieved by implementing fuzzy inference system. The proposed system could translate video signs into text and voice commands. Their data base had 351 gestures with each gesture repeated 10 times by 10 different users. A recognition rate of 96% for gestures in all background environments was achieved. The same authors proposed a complete skeleton of isolated Video Based Indian Sign Language Recognition System (INSLR) [17] that integrates various image processing techniques and computational intelligence techniques in order to deal with sentence recognition.

A wavelet based video segmentation technique was proposed which detects shapes of various hand signs and head movement in video based setup. Shape features of hand gestures were extracted using elliptical Fourier descriptions which to the highest degree reduces the feature vectors for an image. PCA was used to minimize the feature vector again for a particular gesture video and the features were not affected by scaling or rotation of gestures within a video. Features generated using these techniques made the feature vector unique for a particular gesture. Recognition of gestures from the extracted features was done using a Sugeno type fuzzy inference system which used linear output membership functions. Finally the INSLR system employed an audio system to play the recognized gestures along with text output. The system was tested using a data set of 80 words and sentences by 10 different signers. Their system had a recognition rate of 96%. The same authors summarize various algorithms used to design a sign language recognition system [18]. They designed a real time sign language recognition system that could recognize gestures of ISL from videos under different complex backgrounds. They have done a lot of works in the field of ISL recognition. they'd used fuzzy classification and Artificial Neural network classification. Segmenting and trailing of non-rigid hands and head of the signer in signing videos was achieved by pattern active contour models. Active contour energy minimisation was done victimisation signers hand and head color, texture, boundary and form data. Classification of signs was done by associate artificial neural network pattern error back propagation formula. every sign up the video was reborn into a voice and text command. The system has been implemented successfully for 351 signs of Indian language to a lower place all totally different gettable video environments. the popularity rates were calculated for various video environments.

Table 1. Comparison of Various Indian Sign Language recognition Systems

Method	Input	Segmentation	Feature Vector	Classification	Recognition Rate	Platform
Anup Nandy, Soumik Mondal	Real time video	N/A	Orientation Histogram	Euclidean distance	90%	Real time JAVA based software.
Himanshu Lilha, Devashish Shivmurthy	Images	N/A	Histogram of Edge Frequency (HOEF).	Support Vector Machine	98.1%	N/A
P.V.V Kishore, P. Rajesh Kumar	Video	Canny fused Wavelet based	Elliptical Fourier descriptors	Fuzzy	91%	Matlab

P.V.V.Kishore, P.Rajesh Kumar	Real time	Active contours	Texture features	Fuzzy	96%	Matlab 7.0.1
P.V.V.Kishore,P. Rajesh Kumar	Video	Wavelet based	Elliptical Fourier descriptors	Fuzzy	96%	Matlab 7.0.1
P.V.V. Kishore, P.Rajesh Kumar	Real-time	Active contours	Texture features	ANN- error back propagation algorithm	93%	Matlab
Geetha M., Manjusha U C	Images	Boundary tracing	A novel method based on B-spline approximation	Support vector machine	Numbers- 93.2% Alphabets- 91.83%	OpenCV

Geetha M and Manjusha U C proposed [19] a novel vision-based recognition of Indian Sign Language Alphabets and Numerals using B-Spline Approximation. Gestures of ISL alphabets were complex since it involves the gestures of both the hands together. Their algorithm approximates the boundary extracted from the Region of Interest, to a B-Spline curve by taking the Maximum Curvature Points (MCPs) as the Control points. Then the B-Spline curve was subjected to iterations for smoothening resulting in the extraction of Key Maximum Curvature points (KMCPs), which are the key contributors of the gesture shape. Hence a translation & scale invariant feature vector was obtained from the spatial locations of the KMCPs in the 8 Octant Regions of the 2D Space. SVM was used for classification.

V. CONCLUSION

Recent research works have focused mainly on the recognition of static signs of ISL from images or video sequences that have been recorded under controlled

conditions. Special hardware like data gloves, colored gloves or markers is used in some systems. Certain systems use only bare hands but only one background is considered. Most of the systems are signer-dependent and also, the signer needs to wear full sleeve dark colored jackets. Facial features are not included in majority of the systems. Development of signer independent systems which could recognize signs from both facial and hand gestures is a major challenge in the area of sign language recognition. The researchers should focus on the development of a well suited segmentation scheme which is capable of extracting the hand and face region from videos/images having any background. More emphasis should also be given for extracting such features which could completely distinguish each sign irrespective of hand size, distance from the source, color features and lighting conditions.

VI. REFERENCES

[1]. Noor Adnan Ibraheem, RafiqulZaman Khan, 2012, Survey on Various Gesture Recognition Technologies and Techniques, International Journal of Computer Applications, Volume 50, No. 7

[2]. Ping-Sung Liao, Tse-Sheng Chen, Pau-Choo Chung, 2001, A Fast Algorithm for Multilevel Thresholding, Journal of Information Science and Engineering 17, pp. 713-727

[3]. Dr. Alan M McIvor, Background subtraction techniques, Image and Vision Computing Newz Zealand 2000 (IVCNZ00)

[4]. Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai, Skin Segmentation Using Color and Edge Information, Proceedings on International Symposium on Signal Processing and its Applications, 1-4 July 2003, Paris, France

[5]. Jorge Badenas, Josee Miguel Sanchiz, Filiberto Pla, 2001, Motion-based Segmentation and Region Tracking in Image Sequences, Pattern recognition 34, pp. 661-670

[6]. Sanjay Kumar and Dinesh K. Kumar, 2005, Visual Hand Gestures classification Using Wavelet Transform and Moment Based Features, International Journal of Wavelets, Multiresolution and Information Processing, Volume 3, Issue 1

[7]. Qing Chen, Nicolas D. Georganas, Emil M. Petriu, Real-Time Vision-based Hand Gesture Reconition Using Haar-Like Features, Instrumentation and Measurement Technology Conference-IMTC 2007, Parsaw, Poland, May 1-3

[8]. J. Rekha, J. Bhattacharya, and S. Majumder, Shape,Texture and Local Movement Hand Gesture Features for Indian Sign language Recognition, 3rd International Conference on Trendz in Information Sciences and Computing, 8-9 Dec, 2011

[9]. Hyung-Ji Lee, Jae-Ho Chung, Hand Gesture Recognition Using orientation histogram, TENCON 99, Proceedings of the IEEE Region 10 Conference, 1999

[10]. Chieh-Chih Wang and Ko-Chih Wang, Hand Posture Recognition Using Adaboost with SIFT for Human Robot Interaction, Recent Progress in Robotics, LNCIS 370, pp. 317-329, 2008

[11]. Deng-Yuan Huang, Wu-Chih Hu, Sung-Hsiang Chang, Vision-Based hand Gesture Recognition Using PCA+Gabor Filters and SVM, Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009

[12]. Anup Nandy, Soumik Mondal, Jay Shankar Prasad, Pavan Chakraborty and G.C.Nandi, 2010 Recognizing & Interpreting Indian Sign Language Gesture for Human Robot Interaction, International Conf. on Computer & Communication Technology, 2010 |ICCCT'10|, pp. 712-717

[13]. Qing-song Zhu, Yao-qin Xie, Lei Wang (2010) Video Object Segmentation by Fusion of Spatio-Temporal Information Based on Gaussian Mixture Model, Bulletin of advanced technology research, vol. 5, No. 10, pp 38-43.

[14]. P.V.V Kishore, P. Rajesh Kumar, E. Kiran Kumar & S.R.C.Kishore, 2011, Video Audio Interface for Recognizing Gestures of Indian Sign Language, International Journal of Image Processing (IJIP), Volume 5, Issue 4, 2011 pp. 479-503

[15]. Himanshu Lilha, Devashish Shivmurthy, 2011, "Evaluation of Features for Automated Transcription of Dual- Handed Sign Language Alphabets", International Conference on Image Information Processing (ICIIP), 3-5 Nov. 2011

[16]. P.V.V.Kishore, P.Rajesh Kumar, 2012, A Model For Real Time Sign Language Recognition System, International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 6, June 2012 pp. 30-35

[17]. P. V. V. Kishore and P. Rajesh Kumar, 2012, Video Based Indian Sign Language Recognition System (INSLR) Using Wavelet Transform and Fuzzy Logic, IACSIT International Journal of Engineering and Technology, Vol. 4, No. 5, October 2012 pp. 537-542.

[18]. P.V.V.Kishore, P.Rajesh Kumar,2012, Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No.6, pp. 35-47

[19]. Geetha M., Manjusha U. C. 2013,A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation, International Journal of Computer Science and Engineering (IJCSE)

[20]. Tie Yang, Yangsheng Xu, 1994, Hidden Markov Model for Gesture recognition Aseema Sultana, T. Rajapushpa, Vision Based Gesture Recognition for Alphabetical Hand gestures Using the SVM Classifier, International Journal of Computer Science and Engineering Technology, Volume 3, No. 7, 2012

Cite this article as :

Patel Harshil, Patel Jay, Patel Smit, Thakkar Mayur, Dr. Sheshang Degadwala, "Accessible System for Sign Language Computation ", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 2, pp. 222-230, March-April 2021. Available at doi : https://doi.org/10.32628/CSEIT217255
Journal URL : https://ijsrcseit.com/CSEIT217255