

Gesture Acquisition and Recognition of Sign Language

Shubhra Shree², Ashok Kumar Sahoo²

¹Student, Department of Computer Science, Sharda University, India

²Associate Professor, Department of Computer Science, Sharda University, India

Abstract - Differently abled people face a variety of different issue and problems that cut them off from their surroundings. Regardless of all the advancement, we cannot ignore the fact that the conditions provided by the society for the deaf and hard hearing are still far from being perfect. The communication with deaf and hard hearing by means of written text is not as efficient as it might seem at first. This paper discusses sign recognition with particular emphasis on surveying relevant techniques from the areas of recognition approach, problems tackled and hand tracking which can be applied to each task. The main purpose is to help communication between two groups of people, one hearing impaired and one without any hearing disabilities so that the literate deaf and dumb people will get equal position in our society. Sign Language Recognition has become an active area of research nowadays. Existing challenges and future research possibilities are also highlighted.

Keywords: Sign Language, Acquisition, Recognition, Gesture, Hand

I. INTRODUCTION

Reading is requisite for academic achievement and social participation. Deaf and hard hearing children usually lag behind their fellow with normal hearing in reading development. According to a recent study report by the International Disability and Development Consortium [1], at least half of the world's 6.5 crore children with disabilities are kept out of schools because little or no money is budgeted for their needs. Disabled children form a major part of the 12.4 crore kids estimated to be out of school by the United Nation's Out-of-School Children Initiative. Therefore, to cope with this scenario various sign languages have been used so that at least the part of deaf and hard hearing people from the group of differently abled persons can communicate well in the society. At present, sign languages are well known as a natural means for verbal communication of the deaf and hard hearing people. There is no universal sign language, and almost each country has its own national sign language and fingerspelling alphabet. All the sign languages use visual clues for human-to-human communication combining manual gestures with lips articulation and facial mimics. They also possess a specific and simplified grammar that is quite different from that spoken languages. Sign languages are spoken (silently) by a hundred million deaf people all over the world. In total, there are at least 138 living sign languages according to the Ethnologue catalogue, and many of them are national (state)

or official languages of human communication in some countries like the USA, Finland, the Czech Republic, France, the Russian Federation (since 2013) etc [2]. According to the statistics of medical organizations, about 0.1% of the population of any country is absolutely deaf and the most of such people communicate only by sign languages; many people, who were born deaf, even are not able to read. Additionally to conversational sign languages there are also fingerspelling alphabets, which are used to spell words (names, rare words, unknown signs, etc.) letter-by-letter.

Developing algorithms and techniques to correctly recognize a sequence of produced signs and understand their meaning is called sign language recognition (SLR). SLR is a hybrid research area involving pattern recognition, natural language processing, computer vision and linguistics [3]. Sign Language recognition systems can be used as an interface between human being and computer systems. Sign languages are complete natural language with their phonology, morphology, syntax and grammar. A sign language is a visual-gesture language that is developed to facilitate the differently abled persons by creating visual gestures using face, hand, body and arms [4]. Sign language recognition is mainly consisting of three steps: preprocessing, feature extraction and classification. In preprocessing, a hand is detected from sign image or video. In feature extraction, various features are extracted from the image or video to produce the feature vector of the sign. Finally, in the classification, some samples of the images or videos are used for training the classifier then testing the signs in image or video.

There are varied techniques available which can be used for recognition of sign language. Different research scholars have used different techniques according to the nature of sign language and the signs considered. A lot of work has been done on static sign but unfortunately, till date not much research work has been reported for dynamic sign in Indian Sign Language. Aiming to analyze existing technology on the market and under research, we present a brief description of the latest significant features that are the most referenced in the literature.

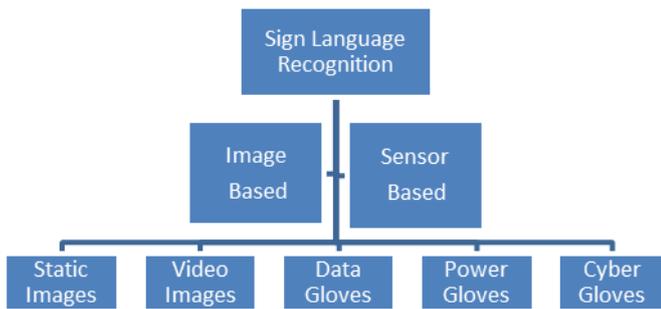


Figure1: Main classification of sign language recognition

II. LITERATURE REVIEW

In this section, the recent work in the area of sign language recognition is discussed. Different researchers use the innumerable types of approaches in recognizing sign language.

In [5], a method for the recognition of 10 two handed Bangla character using normalized cross correlation is proposed by Deb et al. A RGB color model is adopted to select heuristically threshold value for detecting hand regions and template based matching is used for recognition. However, this method does not use any classifier and tested on limited samples. Work on two handed signs has been done in Rekha et al. [6]. Here, Principle Curvature Based Region (PCBR) is used as a shape detector, Wavelet Packet Decomposition (WPD-2) is used to find texture and complexity defects algorithms are used for finding features of finger. The skin color model is used here is YCbCr for segmenting hand region. The classifier used is Multi class non-linear support vector machines (SVM). The accuracy for static signs is 91.3%. However, three dynamic gestures are also considered which uses Dynamic Time Warping (DTW). The feature extracted is the hand motion trajectory forming the feature vector. The accuracy for the same is 86.3%.

In India, research on ISL interpretation started late and very less work is going on at present on ISL continuous word recognition. Kishore and Kumar [7] worked on video based isolated ISL word recognition using fuzzy logic and achieved 96% accuracy. Kalin and Jonas [8] developed educational signing game based on isolated sign recognition of Swedish sign language using Microsoft Kinect. To train the system HMM model was used for a corpus of 51 signs which achieved 89.7% average recognition accuracy. Frank and Sandy [9] used Kinect for interpretation of American sign language for 10 different isolated words. Recognition accuracy of 97% was achieved using support vector machine. Yanhua et al. [10] presented recognition system for Japanese sign language using Microsoft Kinect sensor. A method was developed to employ two Kinects for getting more dataset of hand signs for which point cloud library (PCL) was used to get processed data. Zang et al. [11] used improved SURF algorithm and SVM classifier to recognize

static sign using kinect. Various researchers are working on Arabic sign language recognition for isolated word recognition using various methods such as, pulse coupled neural network (PCNN) [12], HMM [13], simple KNN [14].

Table 1: Comparative study of different approaches

Ref No.	Year	Tracking	Hand features	Approach	Problem tackled	Sign language	Data base Size
15	2013	AAMs ,3D PDM	Geometric measures	HMMs	SL recognition translation	German	~15 K glosses
16	2013	AAMs ,3D PDM	Geometric measures	HMMs	Clustering for sign language analysis	German	~15 K glosses
17	2014	AAMs , 3D PDM	Geometric measures	HMMs	Lip reading in signing	German	~15 K glosses
18	2014	AAMs , 3D PDM	Geometric measures	HMMs	Automatic transcription	British	~10 K glosses
19	2014	Global and local AAM	Shape appearances	Hierarchical clustering	Extreme events detection	Greek American	~10 K glosses
20	2014	Manual annotations	Qualitative relative	ATL, RLDA	Analysis of discriminant non-manual	American	~15 K glosses
21	2016	Randomised	Geometric measures	ATL	Role of counselling and testing	American	~10 K glosses

22	2015	Dyna mic time wrap ping algo ri thm	Geome tric measu res	SVM	Orienta tion of hands	Portu guese	~10 K gloss es
23	2016	AAMs	Shape appear ances	Throu gh analy sis of assis tive techno logies	Potentia l techno logy options	Ameri can	
24	2016	Histo gram s	Segme nt the hand in colour	Using fusion of depth	Complex backgro und	Thai	
25	2015	AAMs	Geome tric measu res	SOA, Nielse n princip les	Signs repre sentations from comput ational semiotic s	Ameri can	
26	2015	AAMs	Geome tric measu res	HMM and k-means	An entropy based k means algo rithm is propose d	Ameri can	
27	2015	AAMs	Geome tric measu res	SVM	Commo n peculiar features	Turki c	
28	2016	Manu al annota tions	Geome tric measu res	Micro soft kinect	Sign languag e with gesture under standing	Russi an	

In [14], dynamic hand gestures having both local and global motions have been recognized through Finite State Machine (FSM). In [15], a methodology based on Transition-Movement Models (TMMs) for large-vocabulary continuous sign language recognition is proposed. TMMs are used to handle the transitions between two adjacent signs in

continuous signing. The transitions are dynamically clustered and segmented; then these extracted parts are used to train the TMMs. The continuous signing is modeled with a sign model followed by a TMM. The recognition is based on a Viterbi search, with a language model, trained sign models and TMM. The large vocabulary sign data of 5113 signs is collected with a sensed glove and a magnetic tracker with 3000 test samples from 750 different sentences. Their system has an average accuracy of 91.9%. Agrawal et al. [16] have proposed a user dependent framework for Indian Sign Language Recognition using redundancy removal from the input video frames. The skin color segmentation and face elimination is performed to segment the hand. Various hand shape, motion and orientation features are used to form a feature vector.

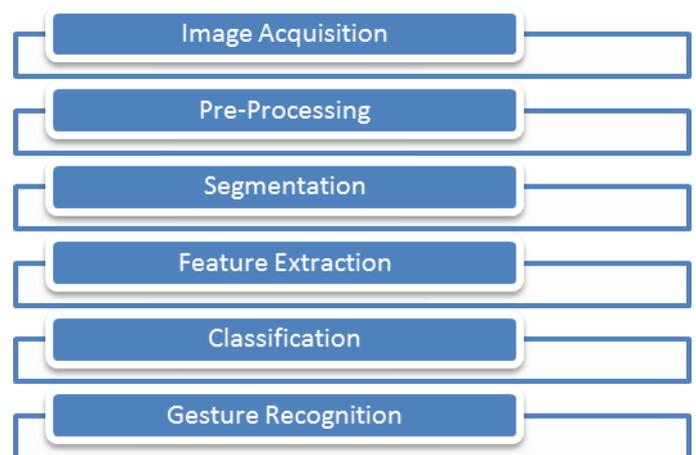


Figure 2: Flow diagram of image base gesture recognition

Another important existing technology is the Leap Motion sensor, a depth sensor made especially to track the hands' features. David Holz, technical director of the Leap Motion company, and Michael Buckwald, co-founder, created a system that allows users to control a digital environment in the same way that objects are controlled in the real world [17].

Existing technologies for this purpose are based on digital image processing and artificial intelligence where are applied techniques and mathematical models able to interpret the captured information [18]. The Microsoft Kinect is currently one of the most used technologies in capturing moving images. However, several technological solutions have emerged, with or without the Kinect, but which are based on capturing images through one or more cameras. Currently, researchers are focusing on adapting the models to three-dimensional scans of the face [19]. Emotion Analysis, developed by Kairos company, offers a facial recognition of emotions and expressions through a simple webcam [20]. The solution that provides Affectiva company is also to be taken into account, offering the Affdex application that analyzes the different facial movements that

can be undertaken and produces the interpretation of emotions from them [21]. Finally a MSVM is used to classify the signs with 95.9% accuracy. For this reason, the use of scanners capable of obtaining high quality 3D images is required. Currently, some companies offer solutions with very positive results, combining some technologies. The Leap Motion controller, associated with the current API, offers positions in Cartesian space of predefined objects, such as fingertips, pen tip, etc. Recognition and interpretation of facial expressions are also fundamental in the sign language recognition.

III. TOOLS FOR GESTURE RECOGNITION

Gesture recognition could be a good example of multidisciplinary analysis. There are totally different tools for gesture recognition, supported the approaches starting from applied math modeling, computer vision and pattern recognition, image process, connectionist systems, etc. Most of the issues are addressed supported applied math modeling, like Principal element Analysis, Hidden Markov Model, Neural Network Classifier and lots of additional advanced particle filtering and condensation algorithms.

While static gesture (hand) recognition will generally be accomplished by template matching, commonplace pattern recognition, and neural networks, the dynamic gesture recognition downside involves the employment of techniques like time-compressing templates, dynamic time deformation, HMM. within the remainder of this section, we tend to discuss the principles and background of a number of these widespread tools utilized in gesture recognition.

IV. SUMMARY OF RESEARCH RESULTS

The following tables show summaries of some hand gesture recognition systems. In this table comparison between recognition strategies in hand gesture recognition strategies used. It provides an outline of application areas and invariant vector of some hand gesture recognition systems. It displays outline of extraction technique, options illustration, and recognition of hand gesture recognition systems that are; hand extraction technique, options vector representation, and recognition employed in the chosen hand gesture recognition systems.

Table 2 : Comparison between recognition methods in hand gesture recognition methods used

Meth od	#Recogniz ed Gestures	#Total Gestures used for Training and Testing	Recogni tion	Database used
[22]	26	1040	DP 98.8%	ASL
[23]	26	208	92.78%	ASL
[24]	0-9 numbers	298 video sequence s	90.45%	Recognize Arabic number from 0 to 9
[25]	5 static/12 dynamic gestures	Totally 240 data are trained and tested	98.3%	5 staic gestures and 12 dynamic
	0-9 numbers	870 training	99.1%	Own database(ISL)

V. CONCLUSION AND DISCUSSION

Comparison analysis of proposed algorithms for dynamic hand gesture recognition revealed that, with any approach, increased vocabulary recognition rate decreases. However, DTW based approach gave better recognition accuracy with more vocabulary than rule based approach. The major advantage of this approach was ISL interpretation system which could interpret meaningful sentence with few input recognized words. However, what was needed was that a sentence was interpreted according to the possible sentence list and keyword that were stored. Sometimes exact sentence might not have been interpreted but thoughts having same meaning were conveyed.

The importance of gesture recognition lies in building economical human-machine interaction. Its applications vary from sign language recognition through medical rehabilitation to virtual reality. during this article, we've provided a survey on gesture recognition, with explicit stress accessible gestures and facial expressions. the main tools surveyed for this purpose include HMMs, particle filtering and condensation rule, FSMs, and ANNs. Plenty of analysis

has been undertaken on sign language recognition, principally victimization the hands (and lips). Facial expression modeling involves the employment of eigen faces, FACS, contour models, wavelets, optical flow, skin colour modeling, as well as a generic, unified feature-extraction-based approach.

A hybridization union of HMMs and FSMs may be a potential study in order to extend the dependability and accuracy of gesture recognition systems. HMMs square measure computationally pricey and need large amount of coaching information. Performance of HMM-based systems may well be restricted by the characteristics of the coaching dataset. On the opposite hand, the statistically prognostic state transition of the FSM would {possibly} possibly cause a lot of reliable recognition. An interesting approach worth exploring is that the freelance modeling of every state of the FSM as associate HMM. this could be helpful in recognizing a fancy gesture consisting a sequence of smaller gestures.

There is no standard dataset available for ISL signs; therefore, we have created our own dataset. Gesture videos were recorded by a digital camera, Sony Cybershot 14 megapixel, placed at 85cm from the subject. A vocabulary of 20 different signs is used for now. The gesture database is divided into training and testing sets. In classification, we have used 4 samples of video of each sign for training and the remaining used for testing. System is trained and tested on multiple signers. The database is composed of varying number of repetitions for each of 22 sign classes which are performed by multiple users. These signers vary in their age ranging from 20-45 years. This framework is user-independent; i.e. the signs trained by one signer can be recognized if different user performs the sign.

REFERENCES

1. <https://www.inshorts.com/news/half-of-worlds-disabled-children-are-out-of-schools-study-1476701217711>
2. Karpov, Alexey, Irina Kipyatkova, and Milos Zelezny. "Automatic Technologies for Processing Spoken Sign Languages." *Procedia Computer Science* 81 (2016): 201-207.
3. 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.
4. T. Dasgupta, S. Shukla, S. Kumar, S. Diwakar, and A. Basu, "A Multilingual Multimedia Indian Sign Language Dictionary Tool," The 6th Workshop on Asian Language Resources, pp. 57-64, 2008.
5. K. Deb, H. P. Mony, and S. Chowdhury, "Two-Handed Sign Language Recognition for Bangla Character Using Normalized Cross Correlation," *Global Journal of Science and Technology*, vol. 12, Issue 3, 2012.
6. J. Rekha, J. Bhattacharya, and S. Majumder, "Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition," in 3rd International Conference onTrendz in information science and computing, pp. 30-35, 2011.
7. P. Kishore,P. Rajesh Kumar,E. Kiran Kumar, and S.Kishore, Video Audio Interface for Recognizing Gestures of Indian Sign Language. *International Journal of Image Processing (IJIP)*,5(4),pp. 479-503, 2011.
8. K. Stefanov, and J.Beskow, A kinect corpus of Swedish sign language Signs. *Proceedings Workshop on Multimodal Corporation*,pp.1-5, 2013.
9. F. Huang, and S. Huang, Interpreting American Sign Language with Kinect pp.1-5, 2011.
10. Y. Sun,N. Kuwahara, and K. Morimoto, Analysis of recognition system of Japanese sign language using 3D image sensor. *IASDR*,pp.1-7, 2013.
11. Z.Hu,L.Yang,L. Luo,Y. Zhang, and X.Zhou, The Research and Application of SURF Algorithm Based on Feature Point Selection Algorithm. *Sensor and Transducers*, IFSA publishing, pp.67-72, 2014.
12. S.Elons, M. Abull-Ela, and M.F.Tolba, A Proposed PCNN Features Quality Optimization Technique for Pose-invariant 3D Arabic Sign Language Recognition. *International Journal of Knowledge-Based Intelligent Engineering System*, vol.14, No.3, pp.139-152, 2010.
13. M. Mohandes,and M. Deriche, Image based Arabic Sign Language Recognition. In *Proceeding 8th Int. Symp. Signal*, vol.1, pp.86-89, 2005.
14. M. K. Bhuyan, D. Ghosh, and P. K. Bora, "A Framework for Hand Gesture Recognition with Applications to Sign Language," *IEEE*, 2006.
15. G. Fang, W. Gao, and D. Zhao, "Large-Vocabulary Continuous Sign Language Recognition Based on Transition-Movement Models," *IEEE Transactions on Systems, Man and Cybernetics, part A*, vol. 37, no. 1, pp. 1-9, 2007.
16. S. C. Agrawal, A. S. Jalal and C. Bhatnagar, "Redundancy removal for isolated gesture in Indian sign language and recognition using multi-class support vector machine," *Int. J. Computational Vision and Robotics*, Vol. 4, pp 23-38, Nos. 1/2, 2014.
17. Leap Motion. Available: <https://www.leapmotion.com/>.
18. Silva VAS. Extração de Atributos para Reconhecimento de Expressões Faciais. 2007.
19. Zhang Z. Microsoft kinect sensor and its effect. *IEEE MultiMedia* 2012;19(2):4-10.
20. Emotion Analysis. Available: <https://www.kairos.com>.
21. Affdex. Available: <http://www.affectiva.com>.
22. Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012 173 Networks," *IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura*.

23. Mokhtar M. Hasan, Pramoud K. Misra, (2011). "Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization", International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2).
24. V. S. Kulkarni, S.D.Lokhande, (2010) "Appearance Based Recognition of American Sign Language Using Gesture Segmentation", International Journal on Computer Science and Engineering (IJCSE), Vol. 2(3), pp. 560-565.