

Indian Sign Language Recognition –A Survey

Mrs.Dipali Rojasara
PG student EC Dept. SVIT,Vasad

Dr.Nehal G Chitaliya
Asst.Professor,EC Dept.SVIT,Vasad

Abstract

Sign language is a mean of communication among the deaf people. Indian sign language is used by deaf or vocally impaired for communication purpose in India. This paper focuses on different techniques used for recognition of Indian Sign language. A review of hand gesture recognition methods for sign language recognition are discussed with different tools and algorithms applied on Indian sign language recognition system, along with challenges & future research direction

Introduction

Sign language is the most natural ways of exchanging information among deaf people. It has been observed that deaf people are facing difficulty to interact with normal people. The purpose of sign language recognition system is to provide an efficient and accurate system to convert sign language into text so that communication between deaf and normal people can be more convenient. Sign language consists of vocabulary of signs in exactly the same way as spoken language consists of a vocabulary of words. Indian sign language (ISL) is sign language used in India. ISL involves both static and dynamic gestures, single as well as double handed gestures, in addition to this the hands involved in gesturing may have complex motion. Some signs include facial expressions too. Because of these difficulties less research work has been done in ISL recognition system [1] A thorough literature survey covering almost all the aspects of the SLR is a primary step to build a ISL recognition system.

The present work reviews a numbers of researches on hand gesture recognition for identifying different signs along with the different steps of the recognition systems. A comparative study of all these works is provided to give

direction to the beginners for their work as well as brief description of the steps associated with sign language recognition system is discussed here.

System Overview

Many researchers have been done on sign language recognition system for different applications, with different recognition phases but they all agree with the main structure of the recognition system. These phases are image acquisition, pre-processing and segmentation, features detection and extraction, and finally the classification or recognition phase. This structure is illustrated in Figure 1.

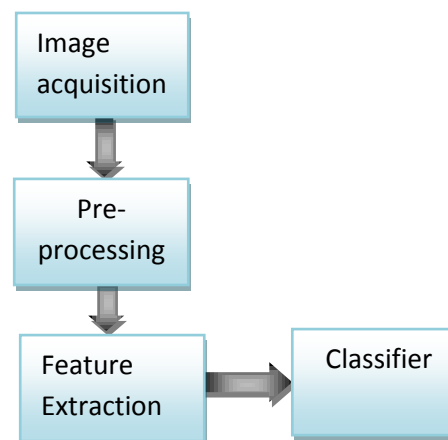


Figure .1 Block Diagram of Sign language recognition system

1. Image Acquisition

The image acquisition of signer, i.e. the person conveying in the sign language, can be obtained by using a camera. The initiation of the acquisition can be done manually. A camera sensor is needed in order to capture the features/ gestures of the signer.

2. Pre-processing

As these images are not taken in a controlled lightening environment also images are taken with a digital camera, they have different sizes and different resolutions. So in image pre-processing is required.

Local changes due to noise and digitization errors should not radically alter the image scene and information. In order to satisfy the memory requirements and the environmental scene conditions, pre-processing of the raw video content is highly important [15]. Various factors like illumination, background, camera parameters, and viewpoint or camera location are used to address complexity of signs. The first most step of pre-processing block is filtering. A moving average or median filter is used to remove the unwanted noise from the acquired image. Background subtraction forms the next major step in the pre-processing block. Running Gaussian average method [16] is used in order to obtain the background subtraction as it is very fast and consumes low memory when compared to other methods. This takes into consideration of the illumination changes like lightning, camera motion changes etc.

CORNELIU LUNGOCI used Image scaling to reduce the computational efforts needed for image processing prior to skin detection. The result of this processing is a binary image in which those pixels that define the hand are colored with white and all the others are black. This processing involves classification of each pixel of the image as part of a human skin or not. There are several techniques developed for skin detection in images, as described in [19]

Segmentation process is necessary for recognizing different signs done with the help of hand. It is the process of dividing the input image (in this case hand gesture image) into regions separated by boundaries[18]. The segmentation process depends on the type of gesture, if it is dynamic gesture then the hand gesture need to be located and tracked [18], if it is static gesture the input image have to be segmented only.

3. Feature Extraction

Features are the crucial elements for sign language recognition. Feature extraction reduces the computational time without sacrificing the accuracy. Large number of features, such as, hand

shape, hand orientation, textures, contour, motion, distance, centre of gravity etc. can be used for sign language recognition. Signs can be recognized using geometric features, like, hand contour, fingertips, finger detections. But these features may neither be always available nor reliable due to occlusions and illuminations [17]. Some non-geometric features (such as colour, silhouette, texture) are also available for recognition. But they are inadequate for the purpose. Therefore, the image or the processed image can be fed to the recognizer to select the features automatically and implicitly, rather than using single type of feature alone. Following approaches are useful for extraction of features.

Principal Component Analysis (PCA) facilitates with advantage of the reduced dimensionality [12]. This technique is used in the work of [12],[11]. However, PCA is very sensitive to the scaling, rotation and translation of the image and hence the image needs to be normalized before applying PCA [12]. PCA works better for the static gesture recognition than dynamic gesture recognition, because normalizing each frame before applying PCA becomes computationally inefficient. DCT used in [1] has the ability to pack most information in few coefficients. It is mainly suited for static sign recognition. Physical features of the shape of the segmented hand contour can be used as feature vectors. This approach has been used in [48], the features related to moment of area, area and perimeter are used for the formation of a feature vector. This approach is simple and computationally fast but has a drawback that it depends on the size of the segmented hand and hence not user independent. The accuracy also reduces with increasing the distance from the camera. So the features extracted are not scale invariant.

Joyeeta Singha et.al. used eigen value & eigen vector for extracting the features. This provides advantages like data compression, data dimension reduction without much loss of information, reducing the original variables into a lower number of orthogonal or non-correlated synthesized variables[20].

4. Classifier

The vision based approach is user friendly as it does not require a signer to wear the cumbersome

gloves. The appearance of the gesture being recognized depends on the position of the camera, distance of the signer from the camera etc. These methods have to maintain a balance between accuracy and computational complexity during the real time performance. The major differences between these two approaches are given in [7].

A. Neural Network

A neural networks can be trained to perform complex tasks of recognition, classification etc. The two types of learning methodologies are explained in [10]. The basic network belonging to unsupervised learning class is Kohonen-Self Organizing Map (Kohonen-SOM). The basics of Kohonen network provided in [6]. Kohonen-SOM was used in [1] for the classification of static gestures comprising some of the alphabets of ISL. The network could give accuracy of 80% for recognition of the static gestures. Two commonly used networks of supervised learning class are Feed Forward Back Propagation Network (BPN) and Radial Basis Function Neural Network (RBFNN). RBFNN was used in [10] for the static gesture recognition of American Sign Language (ASL). Purdue RVL-SLLL ASL database was used for the experiments. The results with this network are very promising with 99% correct recognition accuracy. Feed forward BPN was used in [14] for the classification of static gestures involving alphabets of ASL with recognition accuracy of 92.78%.

B. Finite State Machine

The theory, advantages and algorithm of FSM generation is briefed in [13]. FSM was used in [13] for dynamic and static gesture recognition. The VOPs are extracted from the gesture sequence; subsequently KVOPs are extracted based on based on the Hausdorff distance measure along with their duration, thus creating a FSM. The accuracy obtained for local dynamic gestures is 84.8% and for static gestures it is 97%. FSM was used in [5] for the recognition of both static and dynamic gestures of ASL. The accuracy obtained for dynamic gesture was 61.33% and for static gestures 62.49%.

C. Hidden Markov Model (HMM)

In late 90s the HMM were the primary choice of the gesture recognition technique. A diagrammatic

representation and explanation of the HMM is provided in [10]. HMM was used in [2] for the dynamic gesture recognition involving 60 different dynamic gestures. Recognition was performed using the Viterbi algorithm to estimate maximum likelihood state sequences. The recognition accuracy of 82.17% was achieved. HMM was also used in [5] for recognition of 26 dynamic gestures. The KVOPs are already extracted from the gesture sequence hence Viterbi algorithm was not used, as the algorithm is used to find only the probable sequence. Recognition accuracy of 69.83% was attained.

Conclusion

Main purpose of this paper is to evaluate different technique for recognition of sign language and to provide a comparative study.

The feature extraction facilitates to reduce the computational time without sacrificing the accuracy. For designing a flexible SLR system; the features of an image which are invariant to scaling, translation and rotation are necessary. Depending upon the characteristics of the feature extraction method, single or multiple features can be extracted to achieve an overall invariance.

There is a wide scope for selection of the classifier depending on the speed and the complexity of the desired system. For the recognition of dynamic gestures, HMM and FSM provides robust and faster alternative to the Neural Networks.

References

- [1] Deepika Tewari, Sanjay Kumar Srivastava; "A Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self- Organizing Map Algorithm" ; *International Journal of Engineering and Advanced Technology* , ISSN: 2249 – 8958, Volume-2, Issue-2, December 2012.
- [2] Kenny Morrison, Stephen J. McKenna; "An Experimental Comparison of Trajectory-Based and History-Based Representation for Gesture Recognition"; In *Proceedings of the International Gesture Workshop, 2004*.
- [3] P. Kakumanu, S. Makrogiannis, N. Bourbakis; "A survey of skin-color modeling and detection methods"; *Elsevier, The journal of the pattern recognition society*, 40 , 1106 – 1122, 2007.
- [4] Y. Wang, B. Yuan; "A novel approach for human face detection from color images under complex background", *Pattern Recognition* 34(10) , 1983–1992, 2001.
- [5] Ketki. P. Kshirsagar, Dharmpal Doye; " Object Based key Frame Selection for Hand Gesture recognition";

International Conference on Advances in Recent Technologies in Communication and Computing, 2010.

[6] R. Rojas ; “Neural Networks”; *Springer-Verlag*, Berlin, 1996.

[7] Siddharth S. Rautaray , Anupam Agrawal; “Vision based hand gesture recognition for human computer interaction: a survey”; *Springer Science+Business Media Dordrecht*, November 2012.

[8] Joyeeta Singha, Karen Das; “ Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique”;*International Journal of Advanced Computer Science and Applications*, Vol. 4, No. 2, 2013.

[9] Rashmi D. Kyatanavar, Prof. P. R. Futane; “Comparative Study of Sign Language Recognition Systems”; *International Journal of Scientific and Research Publications*, Volume 2, Issue 6, June 2012.

[10] Anirudh Garg; “Converting American Sign Language To Voice Using RBFNN”; *Master’s Thesis, Computer Science, Faculty of San Diego State University*, Summer 2012.

[11] Bhawna Gautam; “Image Compression Using Discrete Cosine Transform & Discrete Wavelet Transform”; *Master’s Thesis, Computer Science and Engineering, NIT Rourkela*, May 2010.

[12] Henrik Birk, Thomas Baltzer Moeslund; “Recognizing Gestures From the Hand Alphabet Using PrincipalComponent Analysis”; *Master’s Thesis, Laboratory of Image Analysis, Aalborg University, Denmark*, October 1996

[13] M.K. Bhuyan, “FSM-based Recognition of Dynamic Hand Gestures via Gesture Summarization using Key Video Object Planes”, *International Journal of Computer and Communication Engineering* 6, 2012.

[14] Vaishali S. Kulkarni, Dr. S.D.Lokhande; “Appearance Based Recognition of American Sign Language Using Gesture Segmentation”; (*IJCSE*) *International Journal on Computer Science and Engineering*, Vol. 02, No. 03,560-565,2010.

[15] Brian L. Pulito, Raju Damarla, Sunil Nariani, " 2-D Shift Invariant image Classification Neural Network, which overcomes Stability, Plasticity Dilemma", Vol 2, *International Joint Conference on Neural Network*, San Deigo, June 17-21,1990.

[16] Jong Bae Kim,Hye Sun Park,Min Ho Park,Massimo Piccardi,'Background subtraction techniques: a review',*Systems, Man and Cybernetics*,vol.4,*IEEE InternationalConference*,pp:3099-3104,2004.

[17] Murthy, G. R. S. and Jadon, R. S. A Review of Vision Based Hand Gestures Recognition. *Int. J. of Information Technology and Knowledge Management*, 2(2) , 405 – 410 ,2009.

[18] N. Ibraheem, M. Hasan, R. Khan, P. Mishra, “comparative study of skin color based segmentation techniques”, *Aligarh Muslim University, A.M.U., Aligarh, India*,2012.

[19] Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva,” A Survey on Pixel-Based Skin Color Detection Techniques”, *Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University*, 2002.

[20] Joyeeta Singha, Karen Das). “Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique ,”*International*

Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013