

Hand Gesture Recognition for Sign Language Recognition: A Review

Pratibha Pandey¹, Vinay Jain²

¹ Department of Electronics Engineering
Shri Shankracharya Group of Institutions
Bhilai, C.G.

² Department of Electronics Engineering
Shri Shankracharya Group of Institutions
Bhilai, C.G.

Abstract—For deaf and dumb people, Sign language is the only way of communication. With the help of sign language, these physical impaired people express their emotions and thoughts to other person. Since for the common person, it is difficult to understand these language therefore often these physically challenged has to keep the translator along with them to communicate with the world. Hence sign language recognition has become empirical task. Since sign language consist of various movement and gesture of hand therefore the accuracy of sign language depends on the accurate recognition of hand gesture. This paper present various method of hand gesture and sign language recognition proposed in the past by various researchers.

Keywords—SL, DCT, Self organizing Map, Camshift

I. INTRODUCTION

Sign language is one of the effective communication tool for the people who are not able to speak or hear anything. It is also useful for the person who are able to speak but not able to hear or vice versa. Sign language is boon for the deaf and dumb people. Sign language is the combination of different gesture, shape and movement of hand, body and facial expression. Deaf people use these different gesture to represent their thoughts [1]. Each gesture or movement of hand ,facial expression and body movement has special assigned meaning[2]. In different part of the world, different sign languages are used. Sign language of any particular place depends upon the spoken language and

culture of that particular place [1]. For example in USA, American sign language (ASL) is used while in England, british sign language is used(BSL). Similarly Japanese sign language(JSL), French sign language (FSL) and Indian sign language(ISL). Work on Indian sign language has just started after ISL got standardized[3]. This paper presents an extensive review on hand gesture recognition used in sign language recognition which is the heart of sign language recognition.

Aim of developing hand gesture recognition system is to establish an interaction between human and computer and hence recognize the hand gesture automatically which can be used later on for controlling robots or for recognizing the sign language[4][5]. Human –computer interaction(HCI) which is also known as Man-Machine interaction (MMI) is basically represent the relation between human and computer[6][7]. Hand gesture is used for communicating between human and machine as well as between two different people with the help of sign language [8]. Hand gesture may be of static type or dynamic type. In static type of hand gesture, certain position and posture is used for communication purpose. Static type of gesture has less computational complexity [9]. In dynamic type of hand gesture, sequence of posture is used for communication. Since it is dynamic in nature therefore it requires more computational complexity [10]. Dynamic hand gesture is more suitable for real time

application [9][10]. Some of the information is necessary to recognize the hand gesture [11][12]. Some recent review work throw some light on the application of hand gesture recognition in our life [13]. Robot control, Gaming surveillance and sign language recognition are some of the common application of hand gesture recognition[12][14]. The purpose of this paper is to present some review work in the field of hand gesture recognition in the perspective of sign language recognition.

II. HAND GESTURE TECHNOLOGY

Gesture of human body part is one of the important tool for representing the expression and it is considered as one of the natural way to establish a communication between human and computer in virtual environment[15],[16] as well as the between human.



Figure1 Glove based Hand Gesture Recognition



Figure 2 Vision Based Hand Gesture Recognition

Different kind of algorithm is used to capture the gesture and posture of the hand. Vision based approach, Glove based approach and color marker based approach are some of the common techniques used for this very purpose. Figure 1,2 and figure 3 depicts these techniques.



Figure 3 Color Marker Based hand Gesture Recognition

A. Vision Based approach

Vision based approach require camera for capturing the hand gesture or body part gesture this gesture in the form of video or images is then given to the computer for recognition purpose[15]. This approach is most suitable for real time application[3]. Though this approach look simple but actually this approach has many challenges like background problems, variation in lighting condition, color of skin etc. Recognition Time, Computational complexity, robustness are some of the constraint poses by the system. Block diagram of vision based approach used for hand gesture recognition is shown in figure4.

In vision based techniques, bare hand is used to extract the the information required for recognition. The advantage of this method is that it is simple, natural and it directly interact the user with the computer. This technique uses some of the characteristics of images like color, texture for analysing the gesture of hand or other body parts.

Many image processing techniques are used for extracting, segmenting or detecting the hand object from the image. These image processing operation can be divided into two different parts[15] .

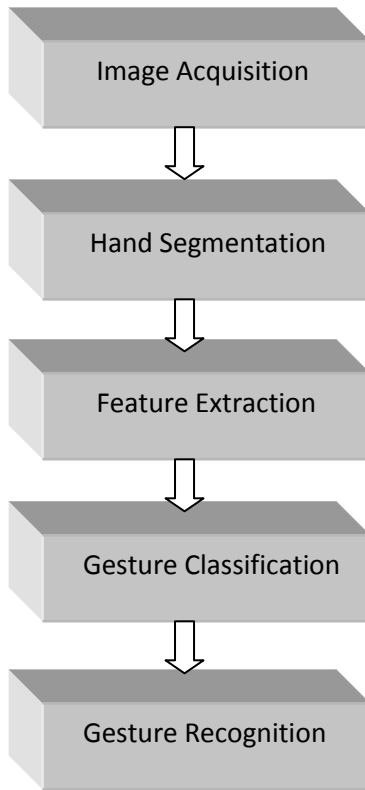


Figure 4 Vision Based Hand Gesture Recognition System

i. Approach based on appearance

In this approach, visual appearance of hand image is modelled with the help of extracted feature of the hand image. These feature are then compared with the feature of the hand image stored in data base. This approach is good for real time application[16] and also it is less complex approach then the 3-D model approach. This is due to the facts that the feature extraction from the 2-d image is easier than the feature extraction from 3-D image. One of the common techniques used in this approach is to detect and segment out the skin color region from the image. This method is affected by the variation in lighting condition, background object with skin like colour. At present a lot of research work is being done to overcome this problems. Adaboost learning algorithm and uses of invariants feature are now being used for extracting out correctly the hand region from the image. This also eliminates the problem of partial occlusion.

ii. Approach Based on 3-D Model

In this approach 3 d model descriptor is used for modelling and analysis of hand shape. In this approach, 2D projection

from 3D model of the hand is made which is used to get kinematic parameters of the hand images. One of the disadvantage of this method is that this approach lead to the loss of some hand features. 3D model can be grouped into two category i.e. Volumetric Model and skeletal model. Volumetric model require huge dimensional parameters while in skeletal model, limited set of parameters are used. Figure 3 represent the 3d Model approach.



Figure 3 3-D model Based Approach for Hand Gesture

Vision based hand gesture recognition require a video camera to acquire the video of hand movements. This input video is then converted in to a frame. Set of predefined features are extracted from each of these frames. Image pre-processing operation is then applied on these frame to make it suitable for segmenting out the hand gesture from the whole frame. Necessary features are extracted from the segmented hand region and stored in the database for training. Extracted features are different for different gesture. Once the feature database is created for all the possible gesture. Then an input video is given to the hand gesture recognition system which convert the video into frames and extract out all the possible feature from the frames. These features are then compared and with the already stored features in the database to interpret the hand gesture.

B. Approach based on Instrumented Glove

In this approach, a glove with attached sensor are worn on the hand. The sensor attached with glove capture the hand movement and position[15]. In this method hand detection is

not required[3]. One of the advantage of this method that it provides accurate position, orientation of the hand, fingers of the palm. The disadvantage of this method that it require the user to connected with the computer physically which make it very uncomfortable technique. This method is also expensive due to the use of sensory gloves.

C. approach based on colored marker.

In this approach, a colored marker hand gloves are worn in the hand. The color of the glove are used to track the movement , position and motion of the hand. This is useful to extract out the geometric feature which is necessary to form hand shape. In color glove, different colors are used to represent different region of the hand such as fingers , palm etc. Advantage of this approach is that it is simple and less expensive as compared to the instrumented glove. Its limitation is that it is also not a natural method of computer human interaction.

III. RELATED WORK

Rajam, P. Subha and Dr G Balakrishnan in their paper[2] presented an algorithm for recognizing the Indian sign language . In their method they have used total 32 signs which represent binary ‘UP’ and ‘DOWN’ position of all five fingers. They developed a 32 combination of binary numbers considering right palm image. In their system, first of all the video is captured depicting particular sign language. This video is then converted in to frames or still images. Then these images are scanned to check the finger tip position of all five fingers. The height of the fingers is measured by taking the bottom of the wrist as a reference position. In order to find the finger coordinate or position, edge detection algorithm is applied to get the edge image of the palm. Upper most point of the edge image is considered as the ‘UP’ position and this edge image is also used to locate the reference point.

Following points are determined in scanning process of the edge image.

- i. Marking of feature points.
- ii. Determining the finger height under ‘UP’ position.
- iii. Determining the angle between the line joining the

feature points of ‘UP’ finger and horizontal line passing through the reference point.

In this method, feature points are computed using left-right scan or Right-Left-Scan mode. All the feature points located by left-right scan mode is marked as ‘Green color’ while the feature points marked by the right-left scan mode is marked as the ‘BLUE color’. Reference points are marked as the ‘Red color’.

Deepika Tewari, Sanjay Kumar Srivastava in their paper[3] presented an algorithm for hand gesture recognition which is further used in Indian sign language recognition system. In their system, self organizing map (SOM) is used for classification purpose. In their approach, first of all the hand image is segmented by utilizing the pixel value for skin and background. Background is represented by the black color. Once the hand image is segmented then discrete cosine transform (DCT) is computed for hand region. 2-d DCT is used for this purpose. Computation of 2-D DCt involves two steps-

- (i) Computing 1-D DCT for each row of the image.
- (ii) Computing 1-D DCT for each column of the image.

After completion of DCT computing, next step is to apply Self Organizing Map (SOM) for classifying the hand gesture and hence the sign language. SOM is a kind of neural network which is also used to classify the data. It is a kind of unsupervised learning. This is advantageous as it doesn’t require manual intervention and this type of learning, little known facts about the input data is sufficient for training.

Kohonen self organizing map is commonly used for this very purpose. In their method, They have used single layer SOM which is feed forward neural network. Since it does not require any prior information about the data therefore it can be used for clustering the data without number of class.

Dhruva N. and Sudhir Rao Rupanagudi [17], suggested in hand gloves worn to recognize the sign language. They have used white color woollen hand gloves for better accuracy. Since in sign language each and every fingers

represent unique message therefore it is very important that each and every fingers must be segmented out clearly. In order to meet this requirement, they have modified the woollen hand gloves. For this purpose each finger in the glove is sown with different color. This makes it easy to segmented out each and every fingers clearly with the help of color based segmentation techniques. RGB and other color based approach is used for segmentation purpose.

Gaolin Fang, Wen Gao, and Debin Zhao in their paper[18] proposed a transition movement model for continuous sign language recognition task. They proposed this method for Chinese sign language recognition. In this method, first of all the sample of sign language is taken as input and fed to the feature extraction module then TMM training and recognition is performed. In the training phase, sign sample of the image is trained to the sign language model. This model then separate the sample into sign parts and transition parts. Temporal clustering algorithm is used to cluster the transition part. This procedure is iterated until the convergence criterion is met. Then in the next phase, with the help of viterbi search algorithm and TMM model is used for recognition purpose.

Sulochana M. Nadgeri, Dr. S.D.Sawarka in their paper[19] suggested Continuous Adaptive mean shift algorithm (CAMSHIFT) tracking algorithm. This algorithm is used to track the color object in video sequence. First of all the color image sequence of the video is represented as probability distribution with the help of color histogram. Since the video sequence is changed during the acquisition time continuously therefore adaptive nature of this algorithm is capable of capturing the dynamics in the sequence and hence track the sequence accurately.

First of all input video is converted in to frames and then CAMSHIFT algorithm is applied to all the frames to find out the frame which covers maximum hand area under the scan window. The tracking is accomplished by following method-

(i) Select the starting location of search window.

(ii) Apply mean-shift operation and store the computed zeroth moment.

(iii) Set the size of search window to the value of zeroth moment obtained in step (ii).

(iv) Repeat step (ii) and (iii) till convergence is met.

Vasiliki E. Kosmidou, and Leontios J. Hadjileontiadis, in their paper [20] suggested electromyogram and 3-D accelerometer based sign language recognition system. In their method, five-channel surface electromyogram and 3-D accelerometer are used to get the data from the dominant hand of signer's. These data is then analysed using intrinsic mode entropy (IMEn) for recognizing the sign language. They have used this method for Greek sign language. Isolated signs are used for recognizing the sign language. Since gesture of the body is directly related to the hand movement therefore measurement of hand movement reflect the gesture of the body. This fact is the motivation behind exploring the possibility of using electro-myogram and 3-D accelerated data for SL recognition. 3-D accelerometer here is used to capture the hand movement while the wrist and finger motion is obtained by corresponding muscles in the arm. With proper experimentation, placement, position and types of sEMG electrode can be determined which gives the optimum signal quality and hence gives the accurate discrimination for different kind of motions.

Joyeeta Singha, Karen Das in their work[21] proposed a Eigen values and Eigen vector based approach for sign language recognition. They have used this method for recognizing Indian sign language. In this approach, first of all, the hand portion up to the wrist is separated out from the hand segmented image and then feature extraction is carried out on this image. Segmented image up to the wrist image is taken and Eigen values and Eigen vector is computed. Once these feature is extracted from all the hand image then classification is carried out. Classification phase consist of two parts-

(i) Performing Euclidean Distance based classification

In the first part Euclidean is computed between Eigen vector of test image and Eigen vectors stored in database. Then we compute the pair which gives minimum Euclidean distance. If Eigen vector of test image is represented by E1 and Eigen vector of database is represented by E2 then the formula is given by

$$ED = \sqrt{\sum_{n=1}^m [E1(n) - E2(n)]^2}$$

(ii) Computing Eigen value weighted Euclidean distance

In this phase, Euclidean distance calculated above is multiplied by the difference between the Eigen values of test image and Eigen values of database image which is stored in variable M as given below-

$$M = E.D. \times |E_{v1} - E_{v2}|$$

Here , E_{v1} and E_{v2} is Eigen value of test image and database respectively.

For each image, sum of the result is computed and added and then minimum of which is considered as the recognized symbol.

Dipak Kumar Ghosh , Samit Ari[22], proposed new algorithm for sign language recognition by applying K-means based Radial basis function neural network. Advantage of this method is that it is rotation invariants for gesture image. In this method first principal component of the hand segmented image is computed and then it is coincided with vertical axis. A feature based on localized contour sequences(LCS) is computed for classifying the hand gesture. Radial basis function neural network based on k-means clustering is used here for classifying the hand gesture from the set of LCS features. The shape of contour very important parameter which can be used to differentiate the hand gesture. Localized contour sequence is basically an effective way of representing the contour and hence is used in this method as feature.

For this purpose in this method first of all, edge image of the static hand need to computed and this job is done by the canny edge detector in this algorithm. Once the edge of the hand contour is computed then contour tracking algorithm is used to find and store the contour pixel. The arrangement of the contour pixel is different for different static hand gesture image. Classification job is accomplished by the radial basis function neural network.

IV. CONCLUSION

This paper deals with the different algorithm and techniques used for recognizing the hand gesture. One of the vital application of hand gesture recognition is to identify the sign language. Sign language is one of the tool of communication for physically impaired, deaf and dumb people. Computer- human interaction and hence computer vision play very important role in this regard by capturing the sign language video and then recognizing the sign language accurately. From the above consideration it is clear that the vision based hand gesture recognition has made remarkable progress in the field of hand gesture recognition. One of the advantage of vision based method is that it can be implemented practically. In order to improve the efficiency of the hand gesture recognition and hence sign language recognition, more work need to be done in the field of feature extraction and classification.

References

- [1] G Adithya V., Vinod P. R., Usha Gopalakrishnan , "Artificial Neural Network Based Method for Indian Sign Language Recognition" , IEEE Conference on Information and Communication Technologies (ICT) ,2013, pp. 1080-1085.
- [2] Rajam, P. Subha and Dr G Balakrishnan , "Real Time Indian Sign Language Recognition System to aid Deaf and Dumb people", 13th International Conference on Communication Technology(ICCT) ,2011, pp 737-742.
- [3] 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 (IJEAT)*, Vol.2, Dec 2012, pp. 165-170.
- [4] G. R. S. Murthy, R. S. Jadon. (2009). “A Review of Vision Based Hand Gestures Recognition,” *International Journal of Information Technology and Knowledge Management*, vol. 2(2), pp. 405-410.
- [5] P. Garg, N. Aggarwal and S. Sofat. (2009). “Vision Based Hand Gesture Recognition,” *World Academy of Science, Engineering and Technology*, Vol. 49, pp. 972-977.
- [6] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, Mo Nours Arab, (2008). “Human-Computer Interaction: Overview on State of the Art”, *International Journal on Smart Sensing and Intelligent Systems*, Vol. 1(1).
- [7] Wikipedia Website.
- [8] 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).
- [9] Xingyan Li.(2003). “Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm”, Department of Computer Science. The University of Tennessee Knoxville.
- [10] S. Mitra, and T. Acharya. (2007). “Gesture Recognition: A Survey” *IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews*, vol. 37 (3), pp. 311- 324, doi: 10.1109/TSMCC.2007.893280.
- [11] Simej 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.
- [12] Joseph J. LaViola Jr., (1999). “A Survey of Hand Posture and Gesture Recognition Techniques and Technology”, Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA.
- [13] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). “Survey on Gesture Recognition for Hand Image Postures”, *International Journal of Computer And Information Science*, Vol. 5(3), Doi: 10.5539/cis.v5n3p110
- [14] Thomas B. Moeslund and Erik Granum, (2001). “A Survey of Computer Vision-Based Human Motion Capture,” *Elsevier, Computer Vision and Image Understanding*, Vol. 81, pp. 231–268.
- [15] Noor Adnan Ibraheem, RafiqulZaman Khan, “Survey on Various Gesture Recognition Technologies and Techniques”,*International Journal of Computer Applications*,Vol. 50, No.7, July 2012, pp. 38-44.
- [16] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, “Vision Based Hand Gesture Recognition”,*World Academy of Science, Engineering and Technology* 49, 2009, pp.972-977.
- [17] Dhruva N. and Sudhir Rao Rupanagudi, Sachin S.K., Sthuthi B., Pavithra R. and Raghavendra , “Novel Segmentation Algorithm for Hand Gesture Recognition”, *IEEE*,2013, pp. 383-388.
- [18] Gaolin Fang, Wen Gao, and Debin Zhao, “Large-Vocabulary Continuous Sign Language Recognition Based on Transition-Movement Models”, *IEEE transactions on systems, man, and cybernetics—part a: systems and humans*, Vol. 37, No. 1, Jan 2007.
- [19] Sulochana M. Nadgeri,Dr. S.D.Sawarkar,Mr.A.D.Gawande, “Hand gesture recognition using CAMSHIFT Algorithm”,*Third International Conference on Emerging Trends in Engineering and Technology,IEEE* ,2010, pp.37-41.
- [20] Vasiliki E. Kosmidou, and Leontios J. Hadjileontiadis, “Sign Language Recognition Using Intrinsic-Mode Sample Entropy on sEMG and Accelerometer Data”, *IEEE transactions on biomedical engineering*, Vol. 56, No. 12, Dec 2009.
- [21] Joyeeta Singha, Karen Das, “Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique”,*(IJACSA) International Journal of Advanced Computer Science and Applications*, Vol. 4, No. 2 , 2013, pp. 188-195.
- [22] Dipak Kumar Ghosh , Samit Ari , “A Static Hand Gesture Recognition Algorithm Using K-Mean Based Radial Basis Function Neural Network”, *IEEE*,2011.