ISSN:0975-9646



# An Efficient Algorithm for Sign Language Recognition

Pratibha Pandey[1], Vinay Jain[2]

[1, 2] *Department of Electronics Engineering*
*Shri Shankracharya Group of Institutions*
*Bhilai, C.G.*

***Abstract*—Sign language is the only tool of communication for the person who is not able to speak and hear anything. Sign language is a boon for the physically challenged people to express their thoughts and emotion. In this paper, a novel sceme of sign language recognition has been proposed for identifying the alphabets in sign language. 10 different spatial domain features ahs been extracted and used for recognition task. Back-propagation neural network is used for classification of features. Results obtained shows a good accuracy of this system in recognizing the alphabets.***

***Keywords*—ASL,ISL,JSL DCT, Self organizing Map, Camshift**

## I. INTRODUCTION

One of the major problems of dumb and deaf people is that they find it difficult to communicate with others. These people require some kind of communication way to communicate with each other and with the others. One of the effective solution of this problem is sign language. In sign language, different gesture of hand is used to express some meaningful information. Using sign language, we can express different alphabet and digit.

Sign language comprises of different gesture, shape and movement of the hand, body and facial expression. With the help of sign language, deaf and dumb people express their different thoughts [1].

Each movement of hand, gesture and facial expression has unique meaning[2]. Language of sign is different in different places. Culture and the spoken language decide the sign language of any particular place.

American sign language(ASL) used in America is different from the british sign language (BSL) of Britain. In the same way Indian sign language(ISL), Japanese Sign language(JSL) and French sign language (FSL) are different from each other .After standardizing the Indian sign language(ISL)[3], work on Indian sign language has started. Recognition of hand gesture is the basis behind developing the efficient sign language recognition system[4][5]. One of the method of hand gesture recognition is to use the hand glove which is used to interact with the computer, populkarly known as the human computer interaction(HCI)[6][7].

With the help of hand gesture, human and machine as well as human to human interaction is possible by using the sign language [8].

There are two types of hand gesture i.e static hand gesture and dynamic hand gesture. In static hand gesture, some pre-defined position and posture is used for communication. Obviously it has less computational complexity [9]. Dynamic hand gesture uses sequences of hand position and posture to represent the sign and hence require more computational power [10][11][12].

In paper [13], some of the application of hand gesture recognition in our daya to day life is explained. Controlling of robot, video gaming and sign language recognition are some of the major application of hand gesture recognition[12][14].

Some recent review work threw 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. Hand gesture recognition can be divided in to three different parts i.e. glove based, vision absed and color marker based hand recognition system[15][16]. In vision based approach camera is required to capture the hand movement. With the help of image processing operation, hand gesture is recognized. In glove based approach, a glove with suitable sensor is worn by the object and different sign language information is captured using the sensor. In color marker based approach, each finger is covered with the different colour. Video is taken and important information is extracted out.


Figure1 Glove based Hand Gesture Recognition

## II. RELATED WORK

Rajam, P. Subha and Dr G Balakrishnan in their paper[2] suggested a Indian sign language recognition system in which UP and DOWN position of all five fingers are used for representing the 32 different sign. 32 different combination of binary numbers are developed in this

system using right palm. In this system edge detection method is used to find out the position and orientation of each finger.

Deepika Tewari, Sanjay Kumar Srivastava[3] suggested a hand gesture recognition algorithm for sign language identification. In this lgorithm first of all the hand part is separated from the background using segmentation algorithm. Twp dimensional DCT is computed for the hand region which acts as the features of the hand. Self organizing map(SOM) is used in this system for classifying the features. Simulation results shows the accuracy of this system which makes this system ideal for using in sign language recognition.

Figure 2 Vision Based Hand Gesture Recognition

Figure 3 Color Marker Based hand Gesture Recognition

In [17] authors Dhruva N. and Sudhir Rao Rupanagudi presented a hand glove based method for hand gesture recognition and hence for sign language rescognition. In their approach, in order to improve the accuracy of the system they have used woollen gloves with different colours for each fingers. This step ensure the accurate segmentation of each fingers by applying colour based segmentation.

Gaolin Fang and Debin Zhao in the paper[18]

Suggested a transition movement model for recognizing the sign language in continuous manner. They prepared their system for Chinese sign language. Video is converted in to a frames then feature extraction method is applied in this algorithm. Training and recognition is performed by the TMM in this project. In the next step this system separate the sign part and transition part. Transition part is clustered by applying the temporal clustering method.

CAMSHIFT (Continuous adaptive mean shift algorithm) trekking based hand gesture recognition algorithm was presented in paper[19]

Electromayogram and 3-D accelerometer can be used for recognizing the hand gesture. 3-D accelerometer is used here to collect the data of hand gesture[20].

Joyeeta Singha, Karen Das in the paper[21]

Presented a sign language recognition system which was based on the eigen values and eigen vector. They have designed this system for recognizing the Indian sign language.

K-mean based radial basis function neural network is used in paper [22] for recognizing the sign language. In this method, principal component analysis is used for feature extraction.

## III. METHODOLOGY

Block diagram of proposed hand segmentation procedure is shown in figure 4.1. In this work first of all, a video is taken for the person signing hand

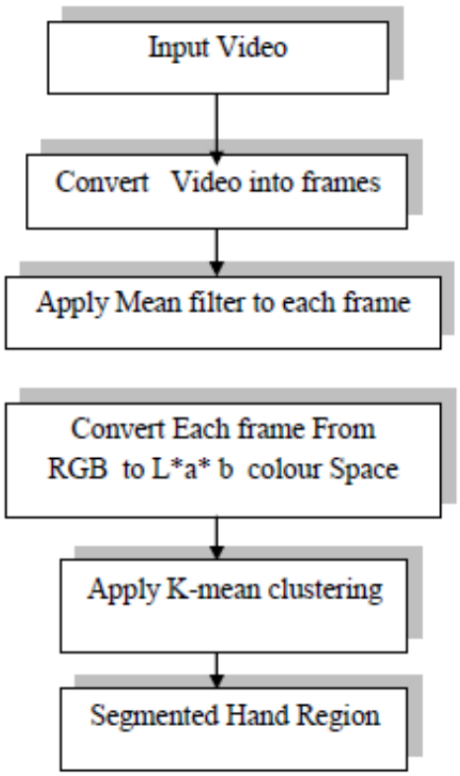



Figure 1 Block Diagram of Proposed method

for different alphabet. Since segmentation of hand region cannot be accomplished directly on the video therefore this video is first of all converted in to frames using Matlab command. Since during the acquisition of video, some noise may add up and causes inaccurate segmentation of hand region. Inappropriate segmentation of hand region may produce false features from the hand therefore , before further processing it is essential to filter out the noise. So in the next step, a mean filter is applied to each and every frame to remove the noise. Once the pre-processing is over

then the next step is to convert the frames from RGB to LAB color space.

During the project work it has been observed that K-means clustering method gives better result in LAB color space therefore in the next step each frame is converted to LAB color space.

So many methods have been proposed in the past for hand segmentation like threshold based, region based, active shape model based approach. Each and every approach has some advantages and disadvantages.

During the project work it has been observed that above mentioned method perform unsatisfactorily in segmenting out the hand region from the image. Apart from this it is also very important that the segmentation method must be clever enough to extract out only the hand region leaving the background un-segmented. Since K-mean clustering based segmentation method is able to overcome above mentioned difficulty therefore it is chosen for segmenting the hand region. In k-means clustering, a user has to input the two or more cluster centre then this method start computing the Euclidean distance of each and every pixel with the cluster centres.

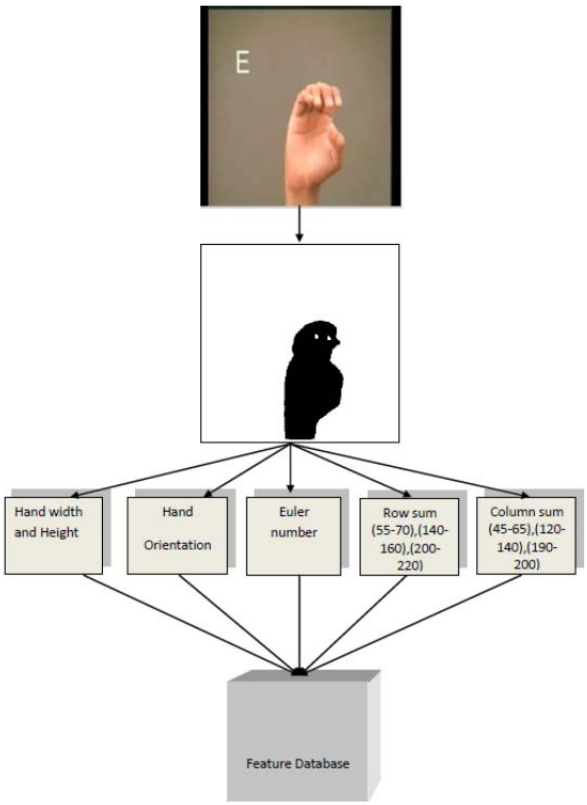
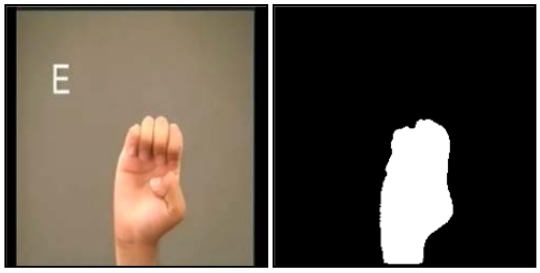



Figure 1.2 Feature Extraction phase

Pixel is assigned to the cluster centre which gives minimum Euclidean distance. In this way the cluster is formed and we get the labelled image. Since in our project we need to segment out hand region from the whole image region therefore two cluster points has been chosen. One for hand region and another for rest of the region. The result obtained by applying this method clearly reveal that this method gives an accurate hand segmented region.




A B

Figure1.3 (a) Actual Frame (B) Segmented Frame

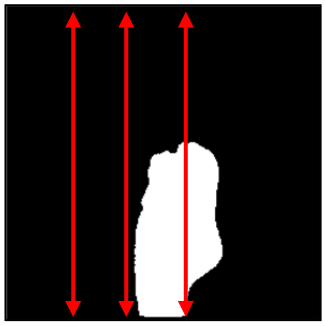


Figure Row sum Feature Extraction

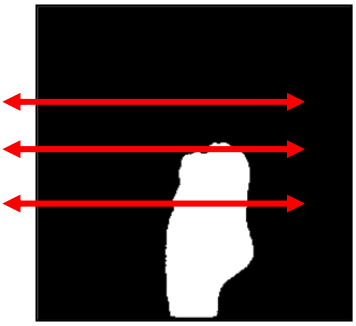


Figure Column Sum Feature Extraction


Figure 1.4 Height of the Segmented image

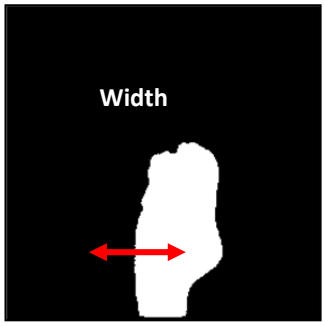



Figure 1.5 Width of the Segmented image

Algorithm steps for HAND Segmentation are summarized as follows-

Step1 - Input Video.
Step2 - Convert Video into a frame(RGB colour space).
Step3 - Apply averaging or mean filter to each and every frame obtained in step2 to remove the noise.
Step4 - Convert each frame fro RGB to L*a*b colour space.
Step5 – Apply K-mean based Segmentation method to segment out hand.

**4.2Feature Extraction**

Once the hand portion is segmented out then the next step is to extract the feature of hands for predicting the hand motion which signify the different sign language. The choice of selecting appropriate feature for hand motion analysis very cumbersome process and it varies from application to application. The features must be appropriate enough to track the hand motion and hence predicting the spoken alphabets. In this project work, as many as 10 different features of segmented hand region have been used. These features are as follows-

1. Width of segmented hand.
2. Height of segmented hand.
3. Orientation of hand.
4. Euler number.
5. Sum of row pixel from row number 55-70.
6. Sum of row pixel from row number 140-160.
7. Sum of row pixel from row number 200-220.
8. Sum of column pixel from column number 45-65.
9. Sum of column pixel from column number 120-140.
10. Sum of column pixel from column number 190-210.

**Width of Hands**= Y-coordinate of right portion of hands- Y-coordinate of left portion of hands.

**Height of Hands** = X-coordinate of lower portion of hands - x-coordinate of upper portion of hands.

**'EulerNumber'** — Scalar that specifies the number of objects in the region minus the number of holes in those objects. (implemented using MATLAB inbuilt command 'regionprops')

**Orientation'** — Scalar; the angle (in degrees ranging from -90 to 90 degrees) between the *x*-axis and the major axis of the ellipse that has the same second-moments as the region. . (implemented using MATLAB inbuilt command 'regionprops'

**Row sum**

$=\sum \#white\ pixel\ in\ a\ selected\ rows\ of\ segmented\ hand$

**Column sum**

$=\sum \#white\ pixel\ in\ a\ selected\ column\ of\ segmented\ hand$

These features have been chosen after extensive research on hand motion. During the research, it has been observed that the width and height of hands changes as per the spoken alphabets.

The variation in orientation and euler number is also different for different alphabets and hence it can also be used for tracking the alphabets. It has also been observed that the Row sum and column sum of particular area is also different for different alphabets.

In order to reduce the dimension of feature vector, instead of taking each feature in each frame, we have taken mean of each feature extracting out from the entire frame. This step reduces the dimension of each feature to great extent. Now each alphabet has only 10 features. Once all the features are extracted from each and every frame of each and every video, then it is stored in database (for each alphabet it is of dimension 10 x number of alphabet sample) which is later used for training the neural network. A target vector is also generated for neural network. the overall database is divided into two part i.e. training database and testing database. In our project 70% of database is used for training and 30% is used for testing.

Once all the features are extracted from the sign language videos of different subjects, the next step is to classify these features for different alphabets. This classification work is accomplished using neural network. In this project work, back propagation neural network is used for classification purpose.

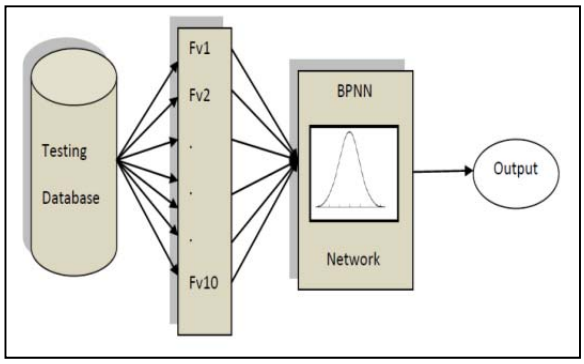



Figure1.6 Training Phase of Neural Network

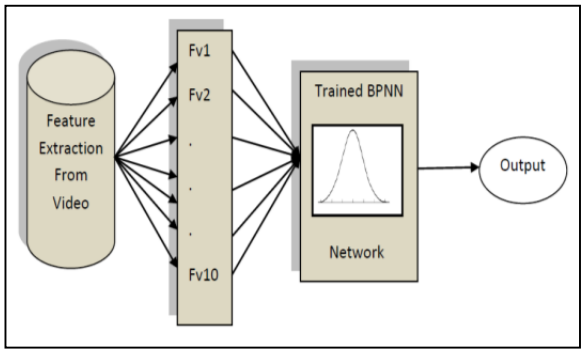



Figure1.7 Testing Phase of Neural Network

Once the input vector and target vector is designed , it is fed to the back-propagation neural network for training purpose. The whole database is divided into two part i.e. training and testing. 70% of the database is set aside for the training and 30% for testing. Once the network is trained then it is tested with the help of testing data.

## IV. EXPERIMENTAL RESULTS

For the proposed method of sign language recognition, a video of sign language of all the 26 alphabet is prepared. For each letter 10 video of different object is taken and video is prepared. So total 260 video has been shot taking different object. A system is designed as per the methodology explained in previous section. First of all a video of each letter is fed to this system for extracting the features of the sign language video. Total 10 different features of each frames are extracted for each alphabet. In each video, middle 10 frames are used for extracting the 10 features. It is done because in each alphabet sign language ,middle frames represent the sign language in best manner. At last mean of all the 10 features extracted from all the middle 10 frames are computed and store in a database. So each alphabet video has 10 features. In this way total 260 features for 26 alphabet has been extracted and store in database. In order to extract the features of the video, video are first converted in to a frames. One of the frames of letter "A" is shown in the figure 6. In order to extract the above mention features from the frames, first of all, in each frames hand part is separated from the background. This is achieved by applying the k means clustering algorithm to each frames. K-means clustering divide the frames in two clustered i.e. Background(Black) and foreground(hand). Figure 7 shows the frame after k means clustering. In this figure some of the hand part become black which is undesirable therefore hole filling operation is applied to figure 7 to get the complete white hand as shown in the figure 8. Figure 9 represent the hand part after segmentation operation. Once the hand part is segmented out then from each hand 10 different features are extracted as described in the methodology part by taking middle 10 frames from all the video. Table 1 shows the 10 different feature value obtained from sign language video of letter "A".

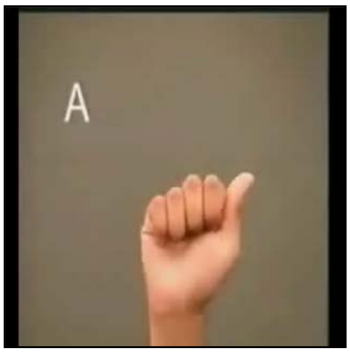



Figure 6 one Frame of letter "A"

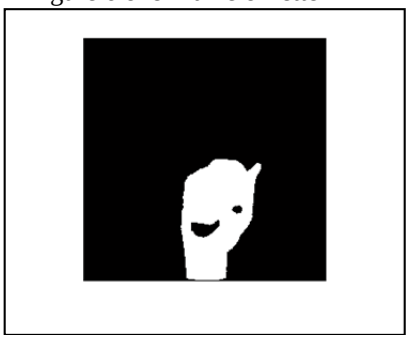


Figure 7 Frame after K mean Clustering

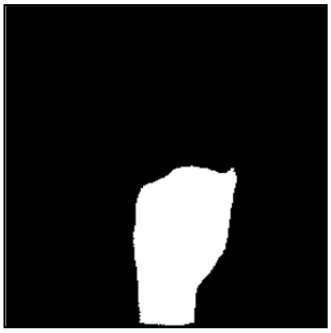


Figure 8 Frame after Hole filling operation

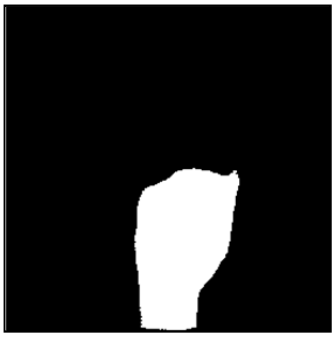


Figure 9 Frame after Segmentation

Table 1 Features for Letter 'A'

| Feature Name | Values |
|---|---|
| Width | 0.1021 |
| Height | 0.1409 |
| Orientation | 0.0663 |
| Euler Number | 0 |
| Row sum1(55:70) | 0.0570 |
| Row sum2(140:160) | 1.4215 |
| Row sum3(200:220) | 1.2241 |
| Column sum1(45:65) | 0.0525 |
| Column sum2(120:140) | 2.2805 |
| Column sum3(190:210) | 0.0211 |

This process is repeated for all the alphabet and then feature database is prepared. Since 10 different videos of each alphabet sign language is taken therefore the dimension of the feature database vector is 10 x 260. In the next step a back-propagation neural network is designed by taking single hidden layer and 26 neurons. Since this type of network is supervised neural network therefore target vector is designed for each alphabet also.

Neural network is then trained from 80 % of feature vector from the feature database.

Rest of the 20% feature vector from the feature database is used for testing the neural network. Neural network is trained and then tested for each alphabet. The output of the neural network for letter "A" is shown in the figure 11. Figure 10 shows the recognition phase of the project.

After testing the proposed method for all the alphabets, recognition accuracy is computed by applying following formula

$$Recognition\ Rate = \frac{Correct\ Match}{Total\ Number\ of\ sample} \times 100$$

Table 5.2 Recognition Rate obtained for different alphabet

| Alphabets | Accuracy (in %) | Alphabets | Accuracy (in %) |
|---|---|---|---|
| A | 95 | N | 97 |
| B | 90 | O | 91 |
| C | 92 | P | 90 |
| D | 95 | Q | 90 |
| E | 87 | R | 92 |
| F | 82 | S | 100 |
| G | 90 | T | 94 |
| H | 91 | U | 94 |
| I | 95 | V | 89 |
| J | 87 | W | 91 |
| K | 96 | X | 92 |
| M | 100 | Y | 92 |
| | | Z | 98 |

Recognition rate obtained after this opeartion is tabulated in table 2.

From this table it is clear that this system recognize the all the letters with more than 90%accuracy .

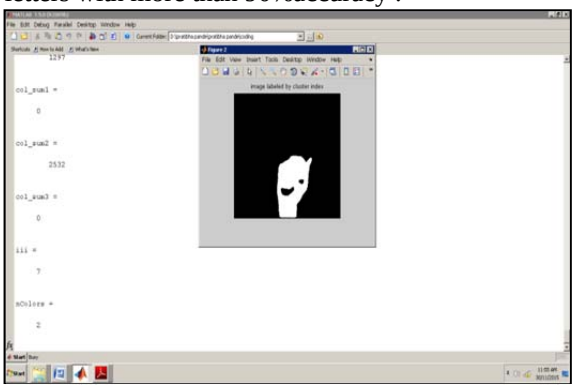


Figure 10 Snapshot of Recognition operation

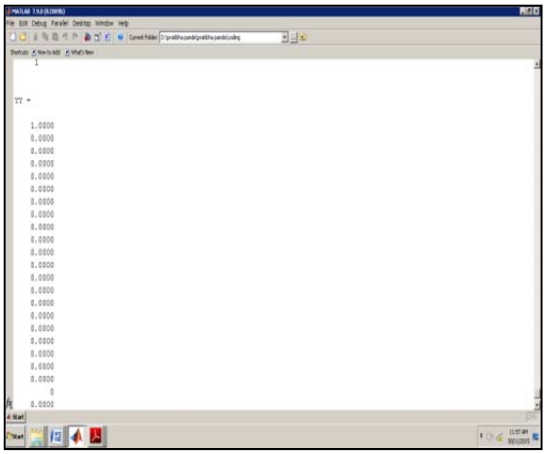


Figure 11 Output of Neural Network for letter 'A'.

## V. CONCLUSION

In this paper an attempt has been made to design a system which can recognize the sign language of alphabets. 10 different features from each video has been extracted to make a feature vector database. neural network is used for classifying the different alphabets and hence for recognition. Accuracy of the proposed method for sign language of different alphabets are tested and found to be more than 90 % for most of the alphabet.

## *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] 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.

[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.