A Novel Hand Segmentation Based Approach For Sign Language Recognition

Swapnil Chandel[1], Akanksha Awasthi[2]
[1, 2] Dept of Electronics & Communication Engg
[2]Asst. Professor, Dept of Electronics & Communication Engg
[1, 2] C.V. Raman University, Bilaspur(C.G.)

Abstract- *sign language based communication is the main tool of correspondence for the individual who can't talk and hear anything. Sign based communication is an aid for the physically tested individuals to express their musings and feeling. In this paper, a novel scheme of sign based communication acknowledgment has been proposed for recognizing the letter sets in communication through signing. 10 distinctive spatial domain features has been extricated and utilized for recognition task. Back-propagation neural system is utilized for feature classification. Results acquired demonstrates a decent accuracy of this framework in perceiving the letter sets.*

Keywords- ASL,ISL,JSL DCT, Self organizing Map, Camshift

I. INTRODUCTION

One of the real issues of dumb and hard of hearing individuals is that they think that its hard to communicate with others. These individuals require some sort of correspondence approach to speak with each other and with the others. One of the viable arrangement of this issue is communication via gestures. In communication through signing, diverse motion of hand is utilized to express some important data. Utilizing gesture based communication, we can express different letters in order and digit.

Communication via gestures involves diverse motion, shape and movement of the hand, body and outward appearance. With the assistance of gesture based communication, not too sharp individuals express their diverse contemplation [1].

Every movement of hand, signal and outward appearance has exceptional meaning[2]. Dialect of sign is distinctive in different places. Culture and the talked dialect choose the gesture based communication of a specific place.

American sign language(ASL) utilized as a part of America is different in relation to the english communication via gestures (BSL) of Britain. Similarly Indian sign language(ISL), Japanese Sign language(JSL) and French communication via gestures (FSL) are not the same as each other .After institutionalizing the Indian sign language(ISL)[3], deal with Indian communication via gestures has begun. Acknowledgment of hand motion is the premise behind building up the proficient communication through signing acknowledgment system[4][5]. One of the strategy for hand motion recognition is to utilize the hand glove which is utilized to interface with the PC, populkarly known as the human computer interaction(HCI)[6][7].

With the assistance of hand motion, human and machine and additionally human to human association is conceivable by utilizing the gesture based communication [8].

There are two kinds of hand signal i.e static hand motion and dynamic hand motion. In static hand signal, some pre-characterized position and stance is utilized for correspondence. Clearly it has less computational many-sided quality [9]. Dynamic hand motion utilizes successions of hand position and stance to speak to the sign and thus require more computational power [10][11][12].

In paper [13], a portion of the use of hand signal recognition in our day to day life is clarified. Controlling of robot, video gaming and communication through signing are a portion of the significant use of hand motion recognition[12][14].

Some current review work tossed some light on the use of hand gesture recognition in our life [13]. Robot control, Gaming observation and communication through signing recognition are a portion of the basic use of hand motion recognition[12][14].

The reason for this section of the paper is to display some survey work in the field of hand motion gesture recognition in the point of view of communication via sign gesture acknowledgment. Hand gesture acknowledgment can be separated in to three distinct parts i.e. glove based, vision based and color marker based hand gesture recognition acknowledgment system[15][16]. In vision based approach camera is required to catch the hand motion. With the

assistance of picture handling activity, hand signal is perceived. In glove based approach, a glove with appropriate sensor is worn by the object and distinctive communication via gestures data is caught utilizing the sensor. In color marker based approach, each finger is secured with the diverse colors. Video is taken and essential data is separated out.

Figure1 Glove based Hand Gesture Recognition

II. RELATED WORK

Rajam, P. Subha and Dr G Balakrishnan in their paper[2] recommended an Indian sign language communication via gestures acknowledgment framework in which UP and DOWN position of each of the five fingers are utilized for speaking to the 32 diverse sign. 32 diverse combination of binary numbers are created in this framework utilizing right palm. In this framework edge recognition strategy is utilized to discover the position and orientation information of each finger.

Deepika Tiwari, Sanjay Kumar Srivastava[3] proposed a hand motion acknowledgment approach for gesture based communication recognizable proof. In this algorithm background is separated from hand captured part with the help of segmentation algorithm. Two dimensional discrete cosine transform I then taken out of the hand portion of the frames. These coefficients work as the features of the hand.

Figure 2 Vision Based Hand Gesture Recognition

Figure 3 Color Marker Based hand Gesture Recognition

These features are then classified with the help of Self organizing map(SOM). The results obtained are very encouraging and shows significant accuracy for this algorithm. So this framework can also be considered as the ideal for sign language recognition.

authors

Dhruva N. and Sudhir Rao Rupanagudi [17] devise a new hand gesture recognition system which was based on the hand gloves. This is very useful method for sign language recognition. This approach utilized the woollen gloves having different colors for each fingers. This is good approach as color based approach can be used for segment out different fingers easily.

Fang and Zhao in their paper [18] proposed transient movement model for recognition of the sign language . This system is designed for the Chinese sign languge. In this system the video of the hand movement or hand gesture is staken. Then this video is converted in to a frames. Features are extracted from each frames. TMM is utilized for training and recognition purpose in this system.

Then sign part and transition part is separated out. All the transition part is separated out by utilizing the temporal based clustering method.

In another approach[19], CAMSHIFT (Continuous adaptive mean shift algorithm) based system is presented for gesture recognition

Electromayogram and 3-D accelerometer based system is alos presented for hand gesture recognition. Feature of the hand motion is collected with the help of 3-D accelerometer[20].

Eigen values and eigen vector can also be utilized for hand gesture recognition. The system based on this is presented by Joyeeta Singha, Karen Das in the paper[21] . this system was designed for Indian sign language. In the paper[22], RBFN neural network which was based on the K-mean is utilized for recognizing the sign language. For feature extraction this system utilizes the principal component analysis(PCA).

III. METHODOLOGY

The methodology for this method actually consist of Two parts

i. Hand Gesture Segmentation From the Video of sign language.
ii. Feature Extraction of the segmented hand Gesture.

For segmenting out the hand gesture from the video frames, an algorithm is designed. The block diagram of this algorithm is shown in the figure 1.

In this approach, first of all, An input video of sign language of different alphabets are taken. In order to segment out the hand gesture from the video sequence, video is converted to the frames with the help of MATLAB internal command.

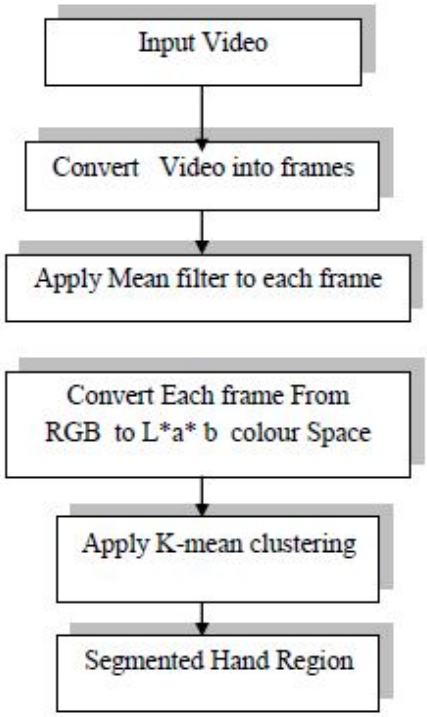


Figure 4 Block Diagram of Proposed method

In order to cancel out the acquisition noise from the video mean filter is applied before any further processing. This step ensure the correct segmentation of the hand from the background. Once the filtering operation is finished, the next step is to segment out the hand region from the background.There are numerous approach are aviaklble for segmenting out the hand from the background like Region based, threshold based, edge based , active contour , Clustering based etc. Among all, the best results are obtained in K-means clustering Based approach.

Therefore segmentation is carried out with the help of K-Mean clustering algorithm. K-mean clustering gives best results For L*a*b color space therefore before applying the k mean clustering the frames is converted in to a Lab color space from the RGB space.

In k-mean Clustering based approach, User gives the two cluster center and then K-mean based approach, compute the Euclidean Distance approach to decide the cluster center to which it belong more.

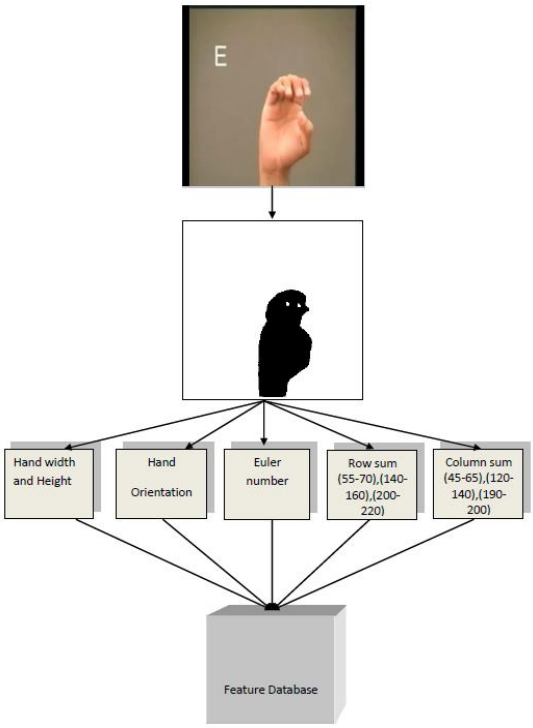
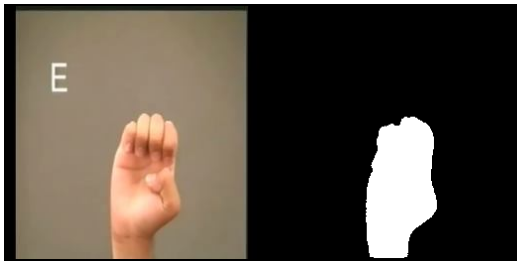


Figure 5 Feature Extraction phase

Euclidean distance between each pixel value and the cluster center is computed and then minimum Euclidean distance is obtained.

The Segmented hand portion is shown in the figure 1.3.



A B

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

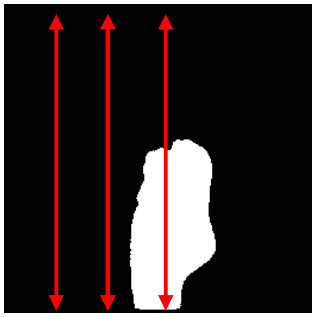


Figure 6 Row sum Feature Extraction

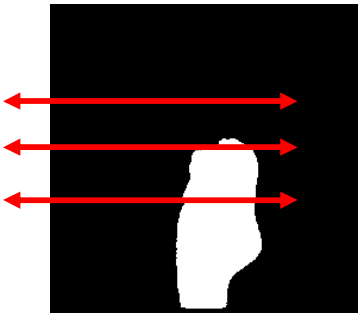


Figure7 Column Sum Feature Extraction

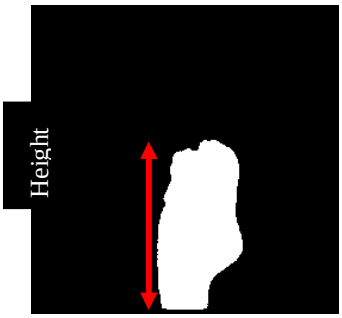


Figure 8 Height of the Segmented image

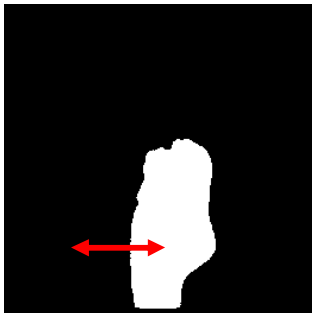


Figure 9 Width of the Segmented image

Below is the summary of the steps of the HAND Segmentation-

Step1 - Input the Sign language Video.
Step2 - Video to Frame Conversion.
Step3- Resize the frames.
Step3 - Application of mean filter in every frames to get rid of noise.
Step4 - Conversion of the RGB to L*a*b color space.
Step5 – Application of K-mean based segmentation approach to segment out the hand portion from the background.

4.2Feature Extraction

Feature extraction is the next step towards the goasl of recognition of sign language. In this phase, some important and unique features of the segmented hand is extracted out. These features are used for predicting the hand motion and hence for signifying the sign language.

Though the choice of selecting the appropriate features is very cumbersome and vary according to the application. After much exploration, as many as 10 different features of the hand region is selected for predicting the sign language. These are as follows-

1. Width of segmented hand region.
2. Height of segmented hand Region.
3. Hand Orientation
4. Euler number of Segmented hand region.
5. Total Sum of row pixel from row number 55-70.
6. Total Sum of row pixel value from row number 140-160.
7. Total Sum of row pixel value from row number 200-220.
8. Total Sum of column pixel value from column number 45-65.
9. Total Sum of column pixel value from column number 120-140.

10. Total Sum of column pixel value from column number 190-210.

Width of Segmented Hand= it is the difference between the Y-coordinate of right portion of segmented hand to the Y-coordinate of left portion of segmented hand.

Height of Segmented Hands = It is the dieffernce between the X-coordinate of lower portion of segmented hand to the x-coordinate of upper portion of segmented hand.

'EulerNumber' — It is a scalar number which gives the difference between number of object to number of holes in those objects. It can be computed by using "regionpropd" command of MATLAB.

'Orientation' — It is also a scalar number i.e angle(from -90 to +90). This is the angle between major axis and the x-axis. Like Euler number, it can also be computed using "regionprops" command of the MATLAB.

Row sum $=\sum$ *#white pixel in a chosen rows of segment hand*

Column sum

$=\sum$ *#white pixel in a chosen column of segment hand*

These features are selected after exploring the hand motion and its relation with different parameters. It has been observed during the research that the height and the width of the hands vary as per the spoken alphabets.

Orientation and the euler number also changes for different alphabets and remain almost intact for similar alphabets. So we can use such parameters for recognition of alphabets. It is also observed that the row sum of pixel and column sum of pixel remain different for different alphabets and almost same for same alphabets.

For reducing the feature vector dimension, mean of each feature is computed for entire frame rather than computing the feature for each frame. This steps ensure only 10 features for each alphabets. These features are stored in the feature database. This feature database is used for training and testing the neural network which is used here for classification of the features. The databse of features are divided into two parts i.e. 70% of the database is set aside for training while the remaining database is used for testing.

In this work back propagation neural network is used for feature classification. Since in such network supervised training is used therefore target vector is also designed for training purpose.

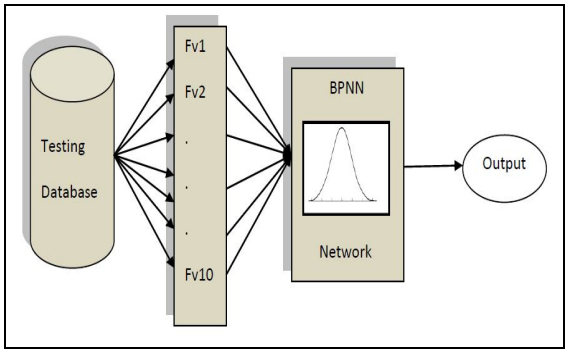


Figure 10 Training Phase of Neural Network

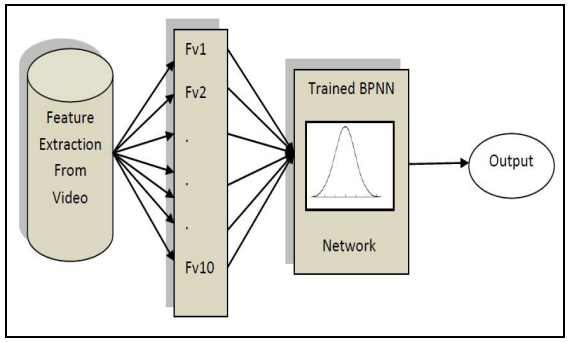


Figure 11 Testing Phase of Neural Network

The training and testing pahse of the neural network is shown in the figure 10 and figure 11.

IV. EXPERIMENTAL RESULTS

A video of sign language for English alphabet are prepared for this project. For each alphabet, 10 video is shot depicting the signlanguage of the alphabets. Therefore total 260 videos are taken. From each video features are extracted as explained in the previous section. Since each video gives the 10 features therefore the dimension of the feature vector becomes 10 x 260. These feature database is then divided in to two parts i.e. 70% is used for training and 30% is used for testing.

For exatracting the features from the alphabet video, first of all video is converted in to a frames. One such frame of letter "A" is shown in the figure12. In this frame, first of all hand part of the image is separated from the background by applying the K-means clustering in L*a*B color space. This method divides the whole frames in to two parts i.e. hand aprt and the background. This is shown in the figure 13. In this figure, the middlae part contain the hole i.e black portion which is undesirable. Therefore hole filling operation is used here for getting rid of such holes. Figure 14 and 15 represent the actual hand segmented part. In the next step, all the 10 features are extracted from this hand part by taking mean of

the middle 10 frames. One such features for alphabet"A" is tabulated in the table 1.

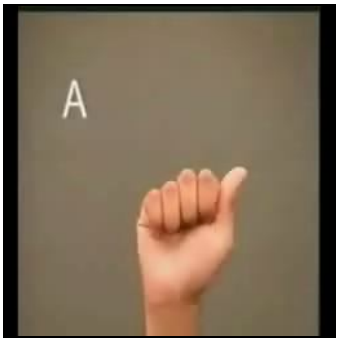


Figure 12 one Frame of letter "A"

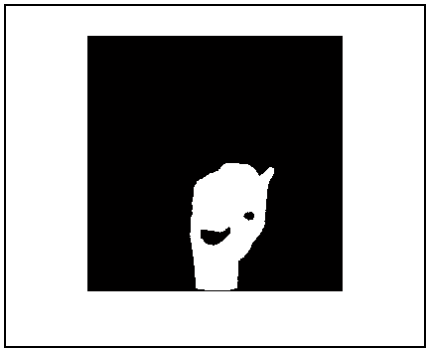


Figure 13 Frame after K mean Clustering

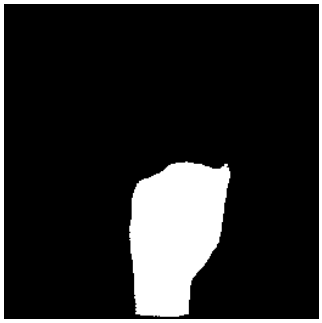


Figure 14 Frame after Hole filling operation

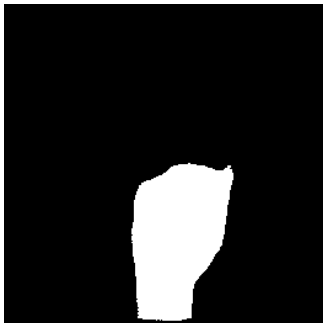


Figure 15 Frame after Segmentation

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 videos and feature database is prepared whose dimension is 10X260. Back propagation neural network is then trained. This neural network has only one hidden layer and 26 neurons.

The output of the neural network for recognizing the letter "A" is shown in the figure 10 and 11.

Recognition accuracy is then computed for different alphabets and it is listed in the table 2

Table 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 accuracy is computed using the below mentioned formula

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

From this table it is clear that the proposed system is capable of producing more than 90% accuracy for all the alphabets

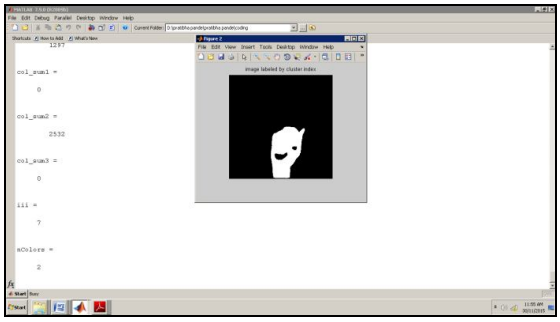


Figure 10 Snapshot of sign recognition

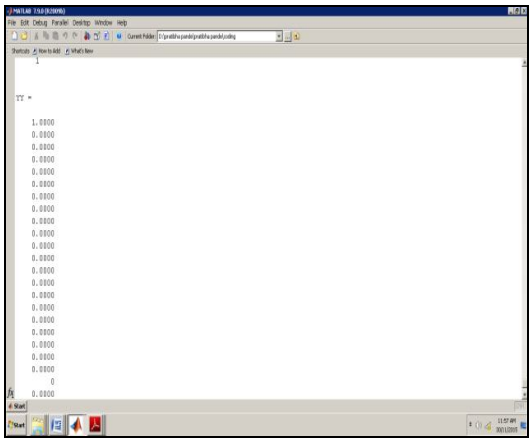


Figure 11 Neural Network output for letter 'A'.

V. CONCLUSION

Designing and implementation of the sign language recognition automatically is very beneficial for the deaf and dumb people. This system helps them to communicate with the other people easily.

This paper presents one such types of system whixh is capable of recognizing the alphabets of sign language video. This system use 10 features and neural network for recognizing the alphabets. The accuracy of the system is found to be of more than 90% for all the alphabets.

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.