International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622
International Conference on Industrial Automation and Computing (ICIAC-12th & 13th April 2014)

RESEARCH ARTICLE OPEN ACCESS

Hand Gesture Recognition of Indian Sign Language to aid Physically impaired People

Neha V. Tavari*, Prof. A. V. Deorankar**, Dr. P. N. Chatur***
*(Department of Computer Science and Engineering,
Government College of Engineering, Amravati, Maharashtra, India)
Email: tavarineha@gmail.com
** (Department of Computer Science and Engineering,
Government College of Engineering, Amravati, Maharashtra, India)
Email: Deorankar.anil@gcoea.ac.in
***(Department of Computer Science and Engineering,
Government College of Engineering, Amravati, Maharashtra, India)
Email: chatur.prashant@gcoea.ac.in

ABSTRACT
Sign language is the language of communication for deaf and dumb people. Most of these physically impaired communities are dependent on sign language translators to express their thoughts to rest of the world. This causes isolation of these people in society. Hence, Sign Language Recognition is one of the most growing fields of research today. A sign language is composed of various gestures formed by physical movement of body parts i.e. hand, arms or facial expressions. In this paper, a method is proposed that makes the use of hand gestures for recognition of Indian Sign Language. Hand Gesture recognition system provides us an innovative, natural, user friendly way of interaction with the computer which is more familiar to the human beings. The proposed method is able to identify the images of the signer which are captured dynamically during testing phase. To implement this approach we have utilized a simple web camera to capture hand gesture images. Artificial neural network is used for recognizing different signs and translate them into text and voice format..

Keywords **-** Hand gesture, Human computer interaction, Physically impaired, Sign language

I. INTRODUCTION

Sign language is widely used by physically impaired people who cannot speak and hear or who can hear but cannot speak and is the only medium of communication for those people. It is nothing but the combination of various gestures formed by different hand shapes, movements and orientations of hands or body, facial expressions and lip-patterns for conveying messages. These gestures are widely used by the deaf and dumb people to express their thoughts [1]. Usually physically impaired people needs the help of sign language interpreters for translating their thoughts to normal people and vice versa . But it becomes very difficult to find a well experienced and educated translator for the sign language every time and everywhere in daily life, but human-computer interaction system for this can be installed anywhere possible. So a system recognizing the sign language gestures automatically is necessary which will help to minimize the gap between deaf people and normal people in the society. The development of a natural input device for creating sign language documents would make such documents more readable for deaf people. Moreover hearing people have difficulties in learning sign language and likewise the majority of those people who were born deaf or who became deaf early in life, have only a limited vocabulary of accordant spoken language of the community in which they live. Hence a system of translating sign language to spoken language would be of great help for deaf as well as for hearing people.

The motivation for developing such helpful application came from the fact that it would prove to be of utmost importance for socially aiding people and it would help increasingly for social awareness as well. Further, if we keep aside this world of computers and just take into consideration human-human interaction, we can realize that we are utilizing a broad range of gesture in personal communication. In fact gesturing is so deeply rooted in our communication that people often continue gesturing when speaking on the telephone. The significant use of gestures in daily life motivates the use of gestural interface in modern era.

As sign language is well structured code gesture, each gesture has a meaning assigned to it [2]. There are number of sign languages spreaded across the world. The sign language used by those deaf and mute at a particular place is dependent on the culture and spoken language at that place.

American Sign. Language (ASL), British Sign Language (BSL), Japanese Sign Language family (Japanese, Taiwanese and Korean Sign Languages), French Sign Language family (French, Italian, Irish, Russian and Dutch Sign Languages), Australian Sign Language, *etc*. [3] are the examples of regionally different sign languages. Indian sign language (ISL) is used by the deaf and dumb community in India and like countries. It consists of both word level gestures and fingerspelling which is used to form words with letter by letter coding. The words for which no signs exist can be expressed with the use of letter by letter signing. It helps in recognizing the words for which the signer does not know the gestures or to emphasis or clarify a particular word. So the fingerspelling has key importance in sign language recognition. ISL differs in the syntax, phonology, morphology and grammar from other country's sign languages. Designing a hand gesture recognition system for ISL is more challenging than other sign languages due to the following reasons.

- Unlike other sign languages (American Sign Language, German Sign language) Indian Sign Language uses both hands to make sign.
- Some signs involve overlapping of both the hands and complicated hand shapes.
- One hand moves faster than the other at times in dynamic hand gestures.

Since ISL got standardized only recently and also since tutorials on ISL gestures were not available until recently, there are very few research work that has happened in ISL recognition [4].Here we propose a method for hand gesture recognition of Indian sign language alphabet and numerals. The signs considered for recognition include 26 letters of the English alphabet and the numerals from 0-9. Indian sign language alphabet and numerals are shown in Fig. 1 and Fig. 2 respectively.

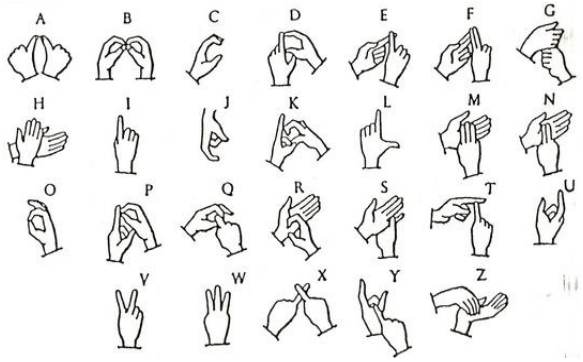

Fig. 1 Representation of ISL Alphabet

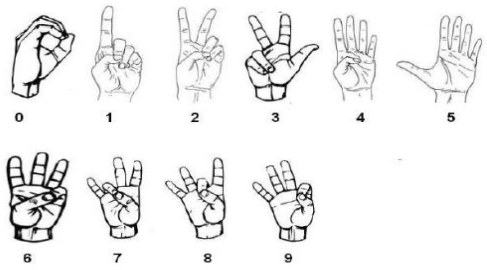

Fig. 2 Representation of ISL Numerals

II. LITERATURE SURVEY

1.1 Hand Gesture Technology

Hand Gestures are considered as the most natural and freer expressive way than other body parts for communications between human and computers in virtual system [5,6] and a powerful means of communication among humans. The key problem in gesture interaction is how to make hand gestures understood by computers [7]. The first step is to collect the data necessary to perform a specific task in order to develop a system. In gesture recognition system different technologies are used for acquiring input data. Present technologies for recognizing gestures can be divided into vision based, instrumented (data) glove, and colored marker approaches. Fig 3 shows an example of these technologies.

Fig. 3 (a) Data-Glove based (b) Vision based (c) Colored Marker[6]

1.1.1 Vision based approaches

In vision based approach, the requirement for system is only the camera(s) to capture the image required for the natural interaction between human and computers without the use of any extra devices. Although these approaches are simple but a lot of gesture challenges are there as these systems need to be background invariant, lighting insensitive, person and camera independent to achieve real time performance. In this technique, a video sequence is captured with the help of camera and the input video is partitioned into frames.A set of features is extracted for each extracted frame sequence. Under image preprocessing the hand object is localized and segmented and the necessary features are extracted and stored in the computer as a trained set. Then each input image pass through the previous steps to extract

its features, and classification algorithms are applied by comparing the extracted features from input image with the training set, to interpret the gesture meaning according to a specific application.

Vision based technologies use a bare hand to extract data needed for recognition , these methods are natural , easy , and the user directly interact with the system. Vision based technology deals with some image characteristics such as texture and color for acquiring data needed for gesture recognition. There are many techniques are applied for detecting hand object after some image preprocessing operations and these methods can be divided into two parts.

i) Appearance based approaches: Appearance based approaches deals with modeling the visual appearance of the input hand gesture image using the feature extracted from the image, which will be compared with the features extracted from stored image . In this approach, there is an advantage of real time performance [7] and easier than 3D model based approaches due to the easier extraction of features in 2D image . The straightforward and common method used in this approach is to identify skin colored regions in the image. Although very popular, but it is very sensitive to changing illumination conditions and other background objects with skin like color. Also there is increased interest in approaches working with local invariant features. The use of invariant features enables the identification of regions or points on a particular object, rather than modeling the entire objects.ii) 3D model based approaches: 3D hand model based approaches rely on 3D kinematic hand model and used to estimate the hand gesture parameters by comparison between the input images and the possible 2D appearance projected by the 3D hand model. But the big disadvantage is that a lot of hand features might be lost in 2D projection .

2.1.2 Instrumented glove approaches

Instrumented data glove approaches use sensor devices for capturing hand gesture position, and motion. Besides the price for these devices are quite expensive, these approaches can easily provide exact coordinates of palm and finger's location and orientation, and hand configurations. These approaches obstacle the ease of interaction between users and computers as it requires the user to be connected with the computer physically.

1.1.2 Colored markers approach

In this approach, the human hand used to wear marked gloves or colored markers with some colors to direct the process of tracking the hand and locating the palm and fingers. It provides the ability to extract geometric features necessary to form hand shape .There might be different colors of color glove, where three different colors are used to represent the fingers and palms, where a wool glove was used. The advantage of this technology is its low cost price comparing with instrumented data glove and simplicity in use. However this technology still limits the naturalness level for human computer interaction to interact with the computer.

1.2 Related Work

Recognition based on 'UP' and 'DOWN' positions of fingers [2] was proposed by a Rajam, P. Subha and Dr G Balakrishnan. In this method, set of 32 signs in which each represents the binary 'UP' & 'DOWN' positions of the five fingers, the most significant bit represents the 'little' finger and the least significant bit represents the 'thumb' finger. A right hand palm images are loaded at run time having 32 combinations of binary number signs. The tip of "UP" fingers is identified by measuring their heights with respect to a reference point at the bottom of palm. The feature points are determined using one of the two scan modes i.e. left-right scan and right-left scan. The feature points located by left-right scan are marked as green in color, those located by Right-Left are marked as blue and reference point is marked as red color.

Deepika Tewari, Sanjay Kumar Srivastava proposed an algorithm for hand gesture recognition system in ISL which is based on vision-based approach. An intensity (grayscale) representation of the segmented image is used for further processing. DCT-based feature vectors are classified to check whether sign mentioned in the input image is "present" or "not present" in the ISL database using self-organizing map (SOM) [3] with unsupervised learning technique in Artificial Neural Network (ANN). As SOM is based on unsupervised learning, no mediation is needed during the learning and little need to be known about the characteristics of the input data which makes it to be used for clustering data without knowing the class memberships of the input data. The SOM is also known as SOFM, the Self-Organizing Feature Map (SOFM) as it can be used to detect features belonging to the problem. The particular kind of SOM known as a Kohonen Network is used which have feed-forward structure with a single computational layer arranged in rows and columns.

The special modified white color woolen hand gloves have been used to simplify the process of gesture identification. Taking into consideration the features of sign language, is shows that each finger in a gesture conveys a particular message and hence

each and every finger has to be identified uniquely. Dhruva N. and Sudhir Rao Rupanagudi, Sachin S.K., Sthuthi B., Pavithra R. and Raghavendra [8] developed a novel segmentation algorithm to meet this requirement in which the woolen hand gloves were modified by replacing and sewing each finger of the glove with a colored cloth for each digit of the hand. Thus, It is utilized a unique color coding for each finger of our hand in order to assist in identifying the fingers. Therefore, segmentation based on various color spaces would be a viable option. In proposed work, two of the most popular color spaces used other than RGB and their conversion from the RGB color space.

i) YCbCr Color Space is a prominent color space used in video and digital photography which consists of Y, Cb and Cr components. Y or luma component is described as the representation of brightness of an image, Cb is described as blue difference chroma component and Cr is described as red difference chroma component.

ii) Hue, Saturation and Intensity (HSI) color space is a color space commonly used in computer graphics applications which represents the points of RGB color space in cylindrical co-ordinates. H or Hue component of color space is described as an optical perception wherein a region seems to be comparable to that of one of the primary colors red, green and blue, S or Saturation described as total amount of purest color distributed throughout the image and I or Intensity is described as the total amount of light leaving through the specific region. With the conversion to any of these above mentioned color spaces it is possible to obtain additional valuable information in order to perform perfect segmentation.

Transition movement models (TMMs) [9] is proposed by Gaolin Fang, Wen Gao, and Debin Zhao to handle transition parts between two adjacent signs in large-vocabulary continuous SLR. For large-vocabulary continuous SLR, TMMs were proposed for continuous Chinese SLR. Sign samples taken from input devices are fed into the feature extraction unit and then input into two related parts i.e.TMM training and recognition based on TMMs.. An approach is made to recognize alphabet characters dynamically from color image sequences using "Continuous Adaptive Mean Shift Algorithm (CAMSHIFT)" tracking algorithm stated in [10] by Sulochana M. Nadgeri, S. D. Sawarkar, A. D. Gawande. The algorithm used here is based on a robust nonparametric technique for climbing density gradients to find the mode(peak) of probability distributions called the mean shift algorithm. Here, it is to find the mode of a color distribution within a video scene. The color image data has to be represented as probability distribution by using color histogram for tracking colored objects in video frame sequences. Hence to deal with dynamically changing color probability distributions derived from video frame sequence, the mean shift algorithm is modified to Continuously Adaptive Mean Shift(CAMSHIFT) algorithm. The tracking accuracy of this algorithm is compared against tolerance to noise, distracters and performance is studied.

Data from five-channel surface electromyogram and 3-D accelerometer from the signer's dominant hand were analyzed by Vasiliki E. Kosmidou, and Leontios J. Hadjileontiadis using intrinsic mode entropy (IMEn) for the automated recognition of Greek sign language (GSL) isolated signs. As the gesture is directly connected to hand movement, measurement of the latter could contribute to the gesture representation in the recognition space. This was the motivation to explore the capabilities of sEMG and 3-D-Acc data in SL recognition. To this end, the movements of the arm can be captured with the help of a 3-D Accelerometer, whereas the motions of the wrist and the fingers can be obtained by the corresponding muscles on the arm. After experimentation on the exact placement and type of the sEMG electrodes, a position that provides with high signal quality and discrimination among the performed motions per channel was identified [11]. Eigen values and Eigen vectors are a part of linear transformations. Eigen vectors are the directions along which the linear transformation acts by stretching, compressing or flipping and Eigen values gives the factor by which the compression or stretching occurs. For recognition of hand gestures, only hand portion till wrist is required, thus the unnecessary part is clipped off using this hand cropping technique. After the desired portion of the image is being cropped, feature extraction phase is carried out. Here, Eigen values and Eigen vectors are found out from the cropped image. Joyeeta Singha, Karen Das [12] have designed a new classification technique that is Eigen value weighted Euclidean distance between Eigen vectors which involved two levels of classification.

i) Classification based on Euclidean Distance

ii)Classification based on Eigen value weighted Euclidean distance:

A novel technique is proposed by Dipak Kumar Ghosh, Samit Ari to obtain a rotation invariant gesture image which coincides the 1st principal component of the segmented hand gestures with vertical axes The shape of the contour is an important property that can be used to distinguish of the static hand gestures from one class to another. The localized contour sequence (LCS), which has been confirmed to be a very efficient representation of contours , is selected as a feature set of the hand

gesture. A contour tracking algorithm is proposed to track the contour of a gesture in the clockwise direction and the contour pixels are numbered sequentially starting from the topmost left contour pixel. After successfully extracting a normalized LCS feature vector of the static hand gesture, The classification job is done via k-mean based radial basis function neural network (RBFNN) [13]. Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V. Pujari stated an efficient algorithm [14] to identify the number of fingers opened in a gesture representing an alphabet of the American Sign Language and introduces a very effective and efficient technique for finger detection.The method has three main phases of processing viz., Edge Detection, Clipping and Boundary Tracing. The first phase having Canny edge operator produces an edge detected image which reduces the number of pixels to be processed at runtime. The second phase clips the undesirable portion of the edge detected image for further processing and the final phase traces the boundary of the image and detects finger tips which aid in finger detection.

III. PROPOSED WORK

The proposed work is aimed to develop a sign language education and recognition platform for hearing impaired peoples and communication system for dumb people to convey their message. The main approaches for analyzing and classifying hand gestures for Human Computer Interaction (HCI) include Glove based techniques and Vision based techniques. The objective of the this work is to build a system that uses natural hand gestures as a modality for recognition in the vision-based setup.

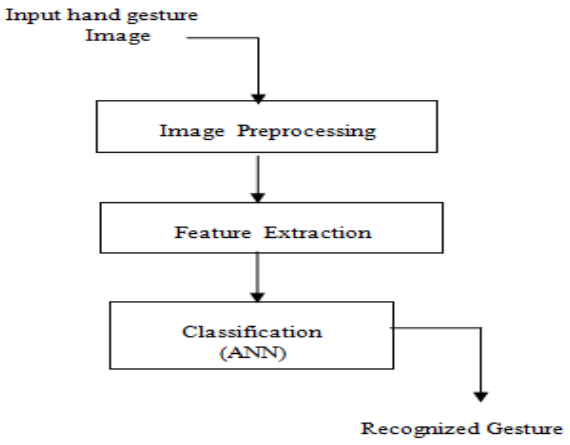

Fig. 4 Block Diagram of Hand Gesture Recognition System

The proposed hand gesture recognition method translates the fingerspelling in Indian sign language to textual and audio form.

- Image Acquisition
- Preprocessing and Segmentation
- Feature Extraction
- Classification

1.3 Image Acquisition

The video sequence of signer, *i.e*. the person conveying in the sign language, can be obtained by using a web camera. Image acquisition is the process to capture the hand gesture images which represents different signs. The image database is created for training and testing phase. As the image dataset of Indian sign language alphabet and numerals are not available from any resources, it is made available with suitable lighting and environmental setup.

Image capturing is a random process. The resolution of various image capturing devices may not be the same. This results in different resolution of the captured images. For accurate comparison of the features and to reduce the computational effort needed for processing, all the images should be scaled to a uniform size [1]. The interface of the application is provided with the button START and STOP. When the user clicks on the first button it works up to open up the integrated webcam and the button changes its status to STOP and when the user is ready with the gesture it can click on the second button so that frame is captured and stored in the directory [5].

1.4 Preprocessing and Segmentation

The image scene and information should not be altered by local changes due to noise and digitization error. Hence to satisfy the environmental scene conditions, preprocessing of the raw data is highly important. The objective of gesture segmentation is to extract the gesture region from the background of the image. Hand segmentation is the process of extracting the hand sign from the captured image and also gesture region is extracted from the background of the image. Efficient hand segmentation has a key role in sign language recognition task. The segmentation process depends on the type of gesture, if it is dynamic gesture then the hand gesture need to be located and tracked , if it is static gesture the input image have to be segmented only [15].

The result of segmentation produces a binary image with the skin pixels in white color and background in black color. The resulting binary image may contain noise and segmentation errors. Filtering and morphological operations are performed on the input image to decrease noise and segmentation errors if any. Image preprocessing includes the set of operations on images whose goal is the improvement of the image data that suppresses undesired distortions or enhances some image features important for further processing.

Fig. 5 Removal of Background

1.5 Feature Extraction

Good segmentation process leads to perfect features extraction process and the later play an important role in a successful recognition process [15]. There are many interesting points on every object which can be extracted to provide a "feature" description of the object. Features vector of the segmented image can be extracted in different ways according to particular application. Under different scene conditions, the performance of different feature detectors will be significantly different. The nature of the background, existence of other objects (occlusion), and illumination must be considered to determine what kind of features can be efficiently and reliably detected. Usually the hand shape and the movement are of major concern in order to recognize gesture. We want gestures to be the same regardless of where they occur with the images boarders. To achieve this we will ignore position altogether, and tabulate a histogram of how often each orientation element occurred in the image. Clearly, this throws out information and some distinct images will be confused by their orientation histograms. Inpractice, however, one can choose a set of training gestures with substantially different orientation histograms from each other.

1.6 Classifiation

Feature vector obtained from the feature extraction step is used as the input of the classifier that recognizes the sign [1]. Training and generalizing are the most basic and important properties of the neural networks. Hence ,Artificial neural network is used as the classification tool. Different network models exist for training the neural net and depending on the feature vectors, the best neural net training method is chosen.

An artificial neural network processes information by creating connections between artificial neurons and they are widely used to model complex relationship between inputs and outputs. Training or learning is used to configure a neural network such that the application of a set of inputs produces a set of desired outputs. Many different algorithms exist to train an artificial neural network. Feedforward backpropagation neural network is used widely for classification purpose. In feedforward neural network, each neuron receives a signal from the neurons in the previous layer, and each of those signals is multiplied by a separate weight value. The weighted inputs are summed, and passed through a limiting function which scales the output to a fixed range of values.The architecture for feedforward network is shown in following figure having x as input and o as output response. The output of the limiter is then broadcast to all of the neurons in the next layer.

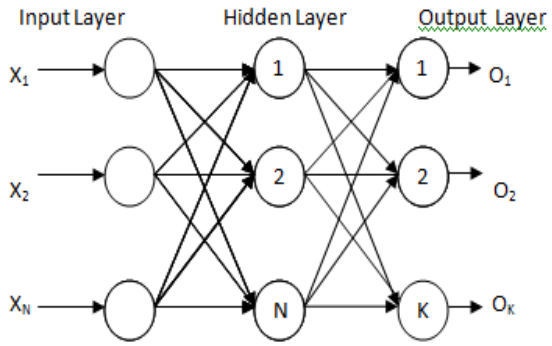

Fig. 7 Feedforward Neural Network

The most commonly used algorithm for training a feedforward neural network is the backpropagation algorithm. The term "backpropagation" describes how this type of neural network is trained. When using a supervised training method, the network must be provided with both sample inputs and anticipated outputs. The anticipated outputs are compared against the actual outputs for given input. Using the anticipated outputs, the backpropagation training algorithm then takes a calculated error and adjusts the weights of the various layers backwards from the output layer to the input layer. Thus, using this algorithm it is possible to recognize each gesture more correctly.

IV. CONCLUSION

A neural network based method for automatically recognizing the hand gestures of Indian sign language is proposed in this work.This paper shows developement in a vision based gesture recognition using a simple system connected with a

web camera. As the method implements completely by using digital image processing technique so the user does not have to wear any special hardware device to get features of hand shape. The features extracted from the sign image are used to train a neural network that recognizes the sign. The system will be able to recognize 36 hand gestures which represents the alphabets from A to Z and numbers from 0 to 9. Developing such system translating sign language to text/voice format will prove very useful for physically imapaired people of India.

REFERENCES

1] 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, Issue 2, Dec 2012, pp. 165-170.

4] Geetha M, Manjusha U C , "A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation", *International Journal on Computer Science and Engineering (IJCSE)*, Vol. 4 ,No. 03,March 2012, pp. 406-415.

5] Sakshi Goyal, Ishita Sharma, Shanu Sharma , "Sign Language Recognition System For Deaf And Dumb People", *International Journal of Engineering Research & Technology (IJERT)*, Vol. 2, Issue 4,April 2013, pp. 382-387.

6] 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.

7] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, "Vision Based Hand Gesture Recognition",*World Academy of Science, Engineering and Technology 49,* 2009, pp. 972-977.

8] 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.

9] 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.

10]Sulochana M. Nadgeri, S.D.Sawarkar, A.D.Gawande, "Hand gesture recognition using CAMSHIFT Algorithm",*Third International Conference on Emerging Trends in Engineering and Technology,IEEE* ,2010, pp.37-41.

11] 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.

12] 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.

13] Dipak Kumar Ghosh , Samit Ari , "A Static Hand Gesture Recognition Algorithm Using K-Mean Based Radial Basis Function Neural Network", *IEEE*,2011.

14] Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V. Pujari, "Finger Detection for Sign Language Recognition",*Proceedings of the international MultiConference of Engineers and Computer Scientists*, Vol. 1, March 2009.

15] Rafiqul Zaman Khan and Noor Adnan Ibraheem, "Hand Gesture Recognition: a Literature Review",*International Journal of Artificial Intelligence & Applications (IJAIA)*, Vol.3, No.4, July 2012, pp. 161-174.