Interpretation of Formal Semantics from Hand Gesture to Text using Proficient Contour Tracing Technique

Sankara Gomathi.S, Amutha .S, Sridhar.G, Jayaprakasan.M

Abstract As an assorted nation with numerous religion tongues, India has attempted to embrace an official, institutionalized gesture based communication. Where as in Indo-Pakistani communication through signing, is viewed as the prevalent sort utilized in South Asia. Many who are deaf or hard of hearing rely on sign language, to communicate. However the estimation of sign language are very unsophisticated and definitions of what counts as proficiency that varies depends on many factors. There are many existing systems which use shape parameters like orientation, palm centroid ,data gloves with 5 accelerometer sensors , and optical markers which reflect infrared light to recognise hand gestures of sign language. Background subtraction techniques used in these systems are K-means clustering ,boundary counters, Eigen backgrounds using Eigen values and wireless technology and bluetooth for connecting software for transmitting recognised hand gesture signals. They are not cost effective but, the accuracy is not met to the need. Whereas, In our proposed system we concentrate mainly to convert hand gestures to text using contour tracing technique to recognise hand gestures using normal webcam. The semantics are classified by support vector machine with trained datasets. The recognised hand gestures are displayed as text. Our main objective is to resolve the problem of facing interviewer for vocally impaired individuals. This helps them to build their confidence and eradicate their inferiority complex compared to other methods.

Keywords: Contour Tracing, Hand gesture, SVM, Feature Extraction, TOF, IoT.

I. INTRODUCTION

The world faces enormous advancement in various technologies in various fields but, there is no appreciable advancement especially for vocally impaired people. They face huge challenges in communicating with the world in their day today life. The sign languages they use is not easily understood by the people at the other end. They cannot share their views conveniently .This leads to the decrease in their confidence level and exposing their talent to the world. Due to lack of confidence they hesitate to face interviews and conference meeting. The system in existence used IoT technology which uses hardware like sensors and gloves. It cannot be expected for the people to have gloves with sensors every time. Due to this limitation, the next phase, implemented system using software. In one the previous approach the hand gesture images were capture by TOF camera and the background was eliminated from the image using background subtraction techniques, and the wrist position was also eliminated from the frame, in order to process the image. The midpoint of palm was found and its maximal radius is obtained. They process the image to convert the grey scale to binary. Finger points were mainly concentrated by using machine learning algorithms to predict the hand gesture with the help of trained data sets. Based on the recognised gesture the text has been displayed. The limitation faced in this existing system was mainly a limited number of hand gesture were processed. The reorganisation of hand gesture were not accurate and thus the result obtained was not correct most of the time as the processing was performed based on finger print section one deals with the Introduction, section 2 is the various literature survey made based on our work, section 3 is our actual proposed work carried out, section 4 is the Algorithms and experimental result, section 5 describes Implementation results and finally we concluded with future work also.

II. LITERATURE SURVEY

The other approach uses contour feature extraction to recognize the hand gestures of particular person . The recognised gesture is converted into text based on the available datasets. The process starts by capturing gesture using web camera. The image is processed using MATLAB library functions. The processed image uses contour tracing and colour segmentation technique through which outline of the palm is obtained. The coordinates set of contours are obtained using contour detection which provides the complete hand outline. Using this outline various properties of contour can be determined. The limitation of this system is less accuracy in user interaction and this has been eliminated in our proposed work.

Another approach for recognising the hand gesture using digital image processing approach, the image is processed using colour segmentation using this technique the skin and non-skin colour objects can be separated then the skin detection process involves in classification of pixels from image by applying the Gary-World algorithm[2]. The sign is

Sankara Gomathi.S, professor& Dean, Department of ECE, Malla Reddy Engineering College and Management Sciences, Affiliated to JNTUH, Telangana, India. (Email: drssgece@gmail.com)

Amutha .S, Assistant Professor, Ramanujan Centre for higher Mathematics, Alagappa University, Karaikudi, Tamil nadu, India.

Sridhar.G, professor. Department of ECE, St.Martin's Engineering college, Telangana, India.

Jayaprakasan.M, Joint Director, Directorate General of Training, MSDE, Government of India, New Delhi, India.

detected using Linear filtering process. The template matching is used to get the target image by comparing trained database. The MATLAB is used to recognise the sign. Linear Discriminates Analysis(LDA) algorithm is used to reduce the dimensionality which provide high accuracy sign recognition.

The other approachis Integrated Two way ISL(Indian Sign Language) .Translation System in which the image is captured using USB camera and then some of the techniques like pre-processing, classification, filtering is used and the image feature vector is obtained. Using the symbol database, the meaning is obtained then they are translated to text or voice.

The advocated system in this paper is to convert sign language to text with software components. The sign language performed by vocally impaired person is captured using normal web camera. The required object from the captured image is recognised with help of particular threshold, that is set prior, while building the system and with the help of contour technique which recognises outline of the hand from background image .The object from image that is recognised is digitally encoded based on representation of the visual characteristics of an object. Now the masked image is obtained, where the background is black in colour and the hand region that needs for processing is white in colour. The noise surrounding the hand region is removed with the help of contour technique. The binarisation is performed to process the image digitally as the content is converted to binary form. The database is loaded with at least 10 different images each representing each signs. The datasets are trained further to improve accuracy for the results of each signs. The classification of each signs in hand gesture is done with the help of convolution neural network technique in deep learning .The pattern for suitable sign in sign language is recognised where pattern matching is performed. The recognised pattern is then compared with trained datasets to obtain the suitable text for the particular sign is performed successfully with Support Vector Machine (SVM). The text is finally displayed to the viewer at the other end.

The novelty proposed in our system is accuracy is achieved as neural network is used for classifying and training datasets. Here the contour technique used to remove unwanted noise that may occur at the background . As our system will be built only with software this can be conveniently used by anyone and anywhere when in need.

III. PROPOSED WORK

This system proposed, in this paper can be used by vocally impaired people to increase their confidence while facing interviews and to attend conference meetings. This can also be used for teaching sign language to young kids who face challenges in communicating. This is also useful for people who are close to those vocally impaired in orderto learn sign language to communicate with them.

The advocated system uses the particular threshold called Otsu threshold, which is one of the binarisation algorithm to convert greyscale to monochrome image in image processing. It includes in rehashing through all the possible limit regards and processing an extent of spread for the pixel levels each side of the edge, for instance the pixels that either fall in frontal region or establishment. The fact of the matter is to find the edge regard, where the entire of frontal region and establishment spreads is at its base. By a pinch of control, we can register what is known as the between class change, which is far speeder to figure. Fortunately, the edge with the most outrageous between class changes also incorporates the base inside class distinction. So it can in like manner be used for finding the best edge and thusly due to being less demanding is an incredibly improved approach to manage use. median filter is used to remove noise from image or signal during the image processing. This filter is basically an non- linear digital filter which helps in preserving certain conditions like edges. Median filter is a smoothing technique to remove noise to smooth areas of signal by providing adverse effect on edges.

A.Pre-Processing operations:

There are three main operation to be carried out to efficacy are listed as below

1.Feature extraction:

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating, how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. However we know that the texture filter functions, described in Texture Analysis cannot provide information about shape, that is, the spatial relationships of pixels in an image.

To create a GLCM, use the graycomatrix function. The graycomatrix function creates a gray-level co-occurrence matrix (GLCM) by calculating how often a pixel with the intensity (gray-level) value i occurs in a specific spatial relationship to a pixel with the value j. By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent), but you can specify other spatial relationships between the two pixels. Each element (i,j) in the resultant GLCM is simply the sum of the number of times that the pixel with value i occurred in the specified spatial relationship to a pixel with value j in the input image are as shown in figure 1.
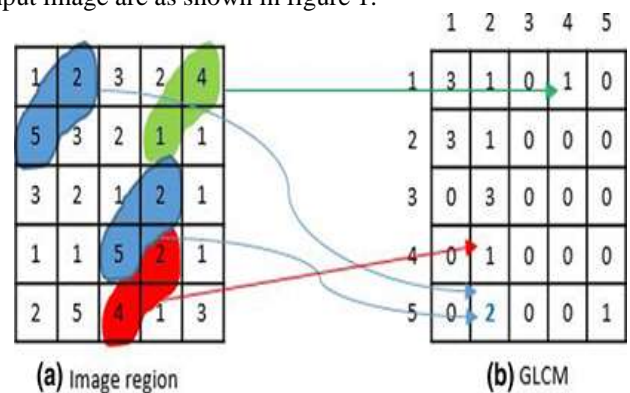

Figure 1: Image region gray level matrix

2. Centroid:

In this development, we process the centroid for distributing swing in to two areas, one which tends to the finger package and other which addresses non finger region. Centroid is also called point of convergence of mass and it portions the swing in to two segments at its geometric fixation provided the picture is uniformly passed on. Centroid is settled utilizing picture minute, which is the weighted run of the mill of pixel's powers of the picture. The centroid is directed by first deciding the picture minute. This is as shown in figure 2.
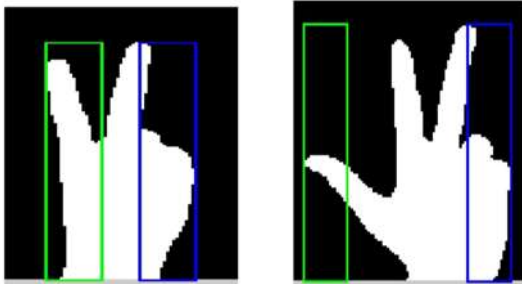

Figure 2: Distribution of centroid thumb recognition

Thumb recognizing confirmation advance is performed popular to see the closeness or nonappearance of thumb close to development. Thumb is considering as an essential shape highlight to portray different hand movements in this framework. We comprehend that thumb can either be stand at right most side of all finger of the hand or at left most side of the pass on general . To perceive the vicinity of thumb close to, we continue with the authoritatively chose swaying box and consider the left side and right 50% of this ricocheting box. By taking 30 pixels width from each side of the jumping box. we crop this skipping encase to two area, one which is tended to by green most remote point is left box and another is correct box tended to by blue limits in the picture appeared as seeks after. Following to having these two boxes we tally the all out number of white pixels appears in twofold picture which address the hand object. Then we check number of white pixels present in every container that is green and blue box. In the event that there is under 7% of all out white pixels exist in any of the right box or left box, we think about that thumb is open in that box in a way. on the off chance that both boxes having over 7% percent of all out white pixels in the picture, by then thumb is absent in any of the compartment .

B. Finger district recognition:

In this development we can note tip of the finger as pinnacle. For getting the rigid number of finger raised close to development, we have to process basically finger locale of the hand that we have in past advancement by dealing with centroid. To continue with this task, we seek after the whole limit structures of hand. Vertical hand picture and level hand picture have been set up in various ways. For vertical hand picture, we just consider the y headings of the purpose of repression frameworks. When we get the estimations of y bearings of purposes of confinement begins stretching out after the sharp decrement in beyond what many would consider possible respect. We consider this sign as tip of the finger and we fix it as a pinnacle respect or a peak. In like way for even picture, we consider the x coordinate of the most remote point cross sections. This times just the x bearings the very pinnacle of lattices is sought after. When

we get the x sort out of purposes of repression begins diminishing after the persistent augmentation we mark this point as a tip of the finger in level hand and set it as pinnacle. Along these lines we found the tip of all brought and fallen fingers up in the picture, in any case we have to bundle massive pinnacles and insignificant peaks among them. For this we have to continue to the going with stage to figure the Euclidean parcel.

1. Euclidean separation:

Subsequent to meaning the perceived apexes or on the other hand tip of the fingers in the hand we ought to find the most amazing best in the hand picture as shown in figure 3. It is with respect to tips of finger and centroid of hand.
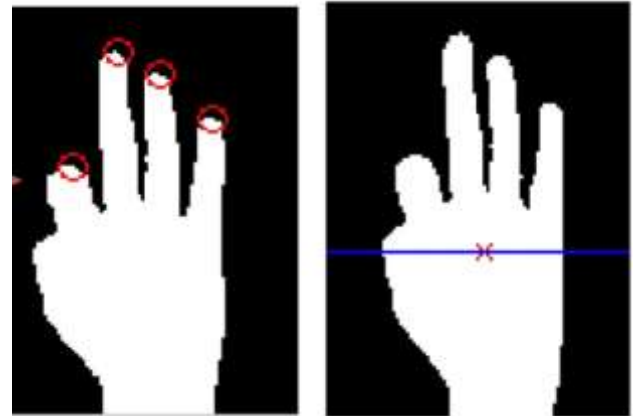

Figure 3: Finger centroid division

2.Peak detection:

Centroid of the division here 'an' addresses all the point of confinement centres and 'b' address the centroid itself. On the reason of this partition formula we can find the length of each raised or fallen finger tolerating centroid as a sort of viewpoint point, this is done as such as to evacuate the watchful number of finger raised in the image. There may be a couple of zeniths recognized which don't everything considered address the tip of the raised fingers, yet the tip of crumbled fingers. We can discard these sorts of insignificant tops by enlisting the most outrageous zenith. Putting the cutoff at 75% of the most outrageous apex regard, we can pick only those essential zeniths whose characteristics are more than this edge a motivating force as these apexes addresses the raised finger near to flag. Distinctive apexes that are distinguished yet don't traverse this edge line would be managed as insignificant zenith or fallen fingers D. Order and bit age Classification of various hand movements relies upon the in this manner made to especially see and utilize these apparent hand motion for supporting human PC affiliation. Zenith Centroid plots are showed up in the basic apexes where we perceived in past development is encoded as '1' and irrelevant zeniths is encoded as '0' in light of the intersection guide status of various fingertips toward farthest point line. Uttermost left piece in the 5 bit twofold gathering is held for status of thumb near to picture. In the occasion that thumbs is accessible, farthest left piece will be 1 by and large 0. This is as shown in figure 4

Figure 4: Hand gestures with bits code and Separation plot.

IV. ALGORITHMS AND RESULTS

1) SUPPORT VECTOR MACHINE

Support Vector Machines (SVM) for explaining design acknowledgment and nonlinear capacity estimation issues have been presented. The likelihood of SVM is mapping the preparation information nonlinearly into a higher-dimensional component space, at that point develop an isolating hyper-plane with most extraordinary edge there. A SVM performs characterization by developing a N-dimensional hyper plane that preferably detaches the information into two classes. SVM models are solidly related to neural system.

A. Stages of Algorithm:

Training stage: preparing tests containing named positive and negative information to the SVM. This information can comprise of separation to fringe vectors, paired pictures, Zernike minutes, and that's just the beginning. Each information is spoken to by vector X_i with name $Y_i = \pm 1$, $1 \leq I \leq l$ is the quantity of models. The choice limit orders all focuses precisely, subsequently $Y_i(wTX_i + b) \geq 1$, for all i variables.

2) CONTOUR:

Contour Tracing known as border following or boundary following. This method is connected to digital pictures so as to remove their limit. Contour following is one of numerous pre-processing systems performed on digital image so as to extricate data about their general shape. At the point when the shape of a given example is separated, it's various characteristics will be examined and used as features which will later on be used in example gathering. Along these lines, right extraction of the form will make continuously exact features which will fabricate the chances of successfully portraying a given example. A "gap seeking" figuring should be used to at first remove the openings in a given example and a short time later apply a shape following count on each gap in order to isolate the absolute form as shown in figure 5.
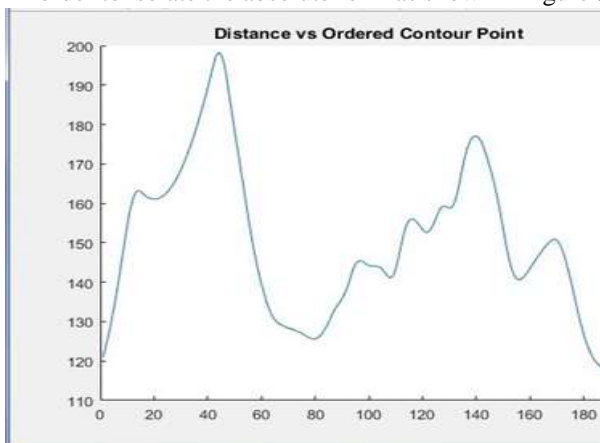

Figure 5: Analysis of distance vs contour point

The system first captures the image using web camera where the hand gesture is shown by vocally impaired person. The Image is segmented such that the region that needs to be processed is extracted from the image by following steps and using algorithm. The obtained image is captured using web camera, has hand gesture that needs to be processed in order to recognize and understand the conversation. This hand gesture is extracted separately from background and noise by applying threshold value .The threshold value is set such that the hand is recognized by HSV colour of skin extracted separately from background and noise by applying threshold value .The threshold value is set such that the hand is recognized by HSV colour of skin .The value ranges from 0 to 255, where 0 is represents as black for background and noise and 255 is represents white. The threshold value is set such that it removes maximum noise from background. Then after threshold is applied the outline of required region is obtained with the help of contour tracing algorithm. Contour tracing algorithm is mainly used to extract the outline of hand region which makes easier to recognize gesture. The pattern for image is performed by classification algorithm, which classifies the signs and recognizes to which sign in trained dataset in database the obtained hand gesture matches. The pattern matching is performed using support vector machine algorithm which provides accurate result from which the recognised symbol is obtained. The text for the recognised symbol is displayed as text in the monitor.

V. IMPLEMENTATION RESULTS

RESIZED IMAGE:

The resized image converts captured image from web camera to pixels for further processing. This is as shown in figure 6.
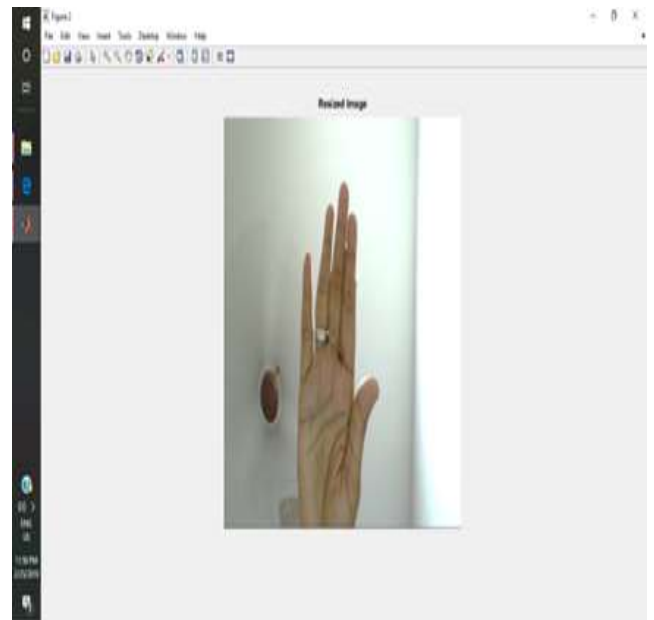

Figure 6: Resized Image

J = imresize(I,scale)

J = imresize(I,[numrows numcols])

[Y,newmap] = imresize(X,map,___)

___ = imresize(___,method)

___ = imresize(___,Name,Value)

J = imresize(I,scale) helps to scale the image and returns the scaled image J.

J = imresize(I,[numrows numcols])The number of rows and columns specified are returned in image J as vectors.

[Y,newmap] = imresize(X,map,___) using colormap map it resizes the indexed image X .

___ = imresize(___,method) This is the method specifies the interpolation method used.

___ = imresize(___,Name,Value) uses name-value pair arguments to control various aspects of the resizing operation.

GRAY SCALE IMAGE:

The Gray Scale Image is a image where intensity of various regions in a image.
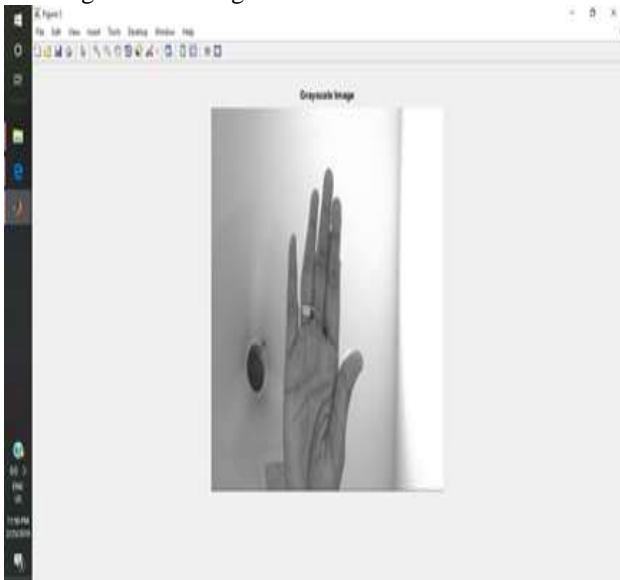

Figure 7: Gray scale image

The RGB coloured image is converted to gray scale image ,using gray scale matrix that helps in obtaining a binary format of image which is black and white (0's and 1's). Matlab makes use of color map for displaying the image is as shown in the above figure 7.

OUTPUT:

The output is obtained in the form of Alert Box by following the above mentioned image processed at various levels as shown in figure 8.
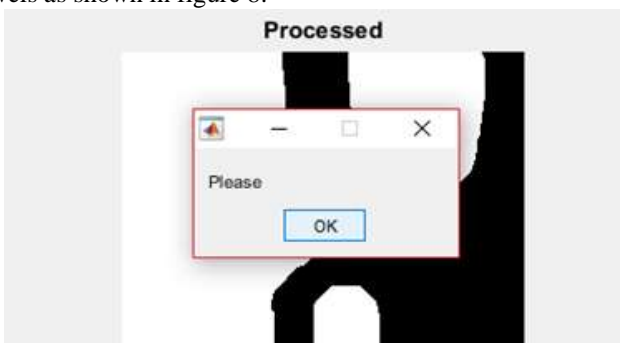

Figure 8. Output of processed image

Productivity is a key proportion of any arrangement demonstrate or speculation. The effectiveness can be estimated in different ways. Be that as it may, this arrangement requires a hit and preliminary strategy as a result of the collaboration between the framework and the object (turn for this situation). The two size of measures is normal yield and the real yield. The normal yield is characterized to be in the scope of 75% to 85% as ideal utilization of calculations is available. The real yield is determined as a proportion of the previously mentioned hit and preliminary strategy.

VI. CONCLUSION

Our paper proposes a framework to assist the deaf and dumb people who face difficulties in communicating with normal person and also people who do not know sign language, without any advanced gadgets like power, data gloves and hued finger top and so on. In the interpretation of framework, conversion of sign to text, Image captured from camera is binaries, noise is expelled, boundaries of finger is detected and corresponding text is displayed as an output to the receiver(an individual who does not know the sign language).We can extend further the same concept to execute in shrewd cell phones too. The challenge in executing this thought in cell phone is actualizing the techniques utilized for picture handling and the accuracy is improved by the results carried.

REFERENCES

1. Rupesh Prajapati "Hand Gesture Recognition And Voice Conversion For Deaf And Dumb" International Research Journal Of Engineering And Technology ISSN 2395-0056 Volume 05 Issue Number 4 April 2018.
2. Mahesh Kumar N B "Conversion of Sign Language into Text" International Journal Of Applied Engineering Research ISSN 0973-4562 volume 13,Issue number 9 2018.
3. Krishna Modi "Translation of Sign Language Finger Spelling To Text Using Image Processing" International Journal Of Computer Application ISSN 0975-8887 Volume 77 ,Issue number 11 November 2013.
4. Kajal B. Bolole,"Gesture recognition using image processing and convertion to text and speech", International journal of exploring emerging trends in engineering, Vol.03,Issue 2, March 2016.
5. Oi Mean Foong, Tan Jung Low, and Satrio Wibowo, " Hand Gesture Recognition: Sign to Voice System S2V"Proceedings Of World Academy Of Science, Engineering And Technology Volume 32 AUGUST 2008 ISSN 2070-3740
6. 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, July 2009.
7. Maebatake, M., Suzuki, I.; Nishida, M.; Horiuchi, Y.; Kuroiwa, S. " Sign Language Recognition Based on Position and Movement Using Multi-Stream HMM", Second International Symposium on, ISBN: 978-0-7695-3433-6.

8. Anuja V.Nair, Bindu.V, "A Review on Indian Sign Language Recognition", International journal of computer applications, Vol. 73, pp: 22, (2013).
9. J. Rekha, J. Bhattacharya, and S. Majumder, "Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition", IEEE 3 rd International Conference on Trendz in Information Sciences & Computing (TISC2011).
10. Rajam, P. Subha and Dr G Bala krishnan, "Real Time Indian Sign Language Recognition System to aid Deaf and Dumb people", 13thInternational Conference on Communication Technology (ICCT).
11. Meenakshi Panwar,"Hand Gesture Recognition based on Shape Parameters" International Conference on Computing, Communication and Application (ICCCA), pp: I-6, IEEE, (2012).
12. Paulraj M.P.Sazali Yaacob,Modh Shuhanaz bin Zanar Azalan,Rajkumar Palaniappan," A Phoneme based sign language recognition system using skin color segmentation",Signal Processing and its Application (CSPA)-pp:1-5,2010.
13. Manar Maraqa,Dr.Raed Abu Zaiter ,"Recognition of Arabic sign language using recurrent neural networks",Applications of International journal of Scientific &Engineering Research,Volume 6 Issue 9 ,September 2015 ISSN 2229-5518.
14. Dr.Sami M.Halawani , " Arabic Sign Language Translation System on Mobile Devices",UCSNS International Journal of Computer Science and Network Security ,Vol 8,no 1,pp.251-256,January 2008.
15. Hee-Deok Yang ,Stan Sclaroff and Seong – WhanLee,"Sign Language Spotting with a Threshold Model Based on Conditional Random Fields",IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol.31,no7 ,pp 1264-1277 , july 2009.
16. Omar AI-Jarrah ,Alaa Halawani ,"Recognition of gestures in Arabic Language using neuro-fuzzy systems" Artificial Intelligence :Else Vier Artificial Intelligence –AI ,vol 133,no 1-2pp.117-138,2001.