# Interpretation of sign language into English using NLP techniques

**Article** · August 2017

DOI: 10.1080/02522667.2017.1372136



**2 authors:**

Sampada Wazalwar
G. H. Raisoni College of Engineering, Nagpur
**3** PUBLICATIONS **1** CITATION
SEE PROFILE

Urmila Shrawankar
G H Raisoni College of Engineering, Nagpur, (MS) India
**162** PUBLICATIONS **281** CITATIONS
SEE PROFILE


Taylor & Francis
Taylor & Francis Group

Journal of Information and Optimization Sciences

ISSN: 0252-2667 (Print) 2169-0103 (Online) Journal homepage: http://www.tandfonline.com/loi/tios20

# Interpretation of sign language into English using NLP techniques

Sampada S. Wazalwar & Urmila Shrawankar

**To cite this article:** Sampada S. Wazalwar & Urmila Shrawankar (2017) Interpretation of sign language into English using NLP techniques, Journal of Information and Optimization Sciences, 38:6, 895-910, DOI: 10.1080/02522667.2017.1372136

**To link to this article:** http://dx.doi.org/10.1080/02522667.2017.1372136

*Journal of Information & Optimization Sciences*
Vol. 38 (2017), No. 6, pp. 895–910
DOI : 10.1080/02522667.2017.1372136



# Interpretation of sign language into English using NLP techniques

Sampada S. Wazalwar *
Urmila Shrawankar
*G. H. Raisoni College of Engineering*
*Nagpur 440016*
*Maharashtra*
*India*



**Abstract**

Sign language is a way of communication for deaf & dumb. Different sign recognition techniques are there which are giving output in the form of word for recognized sign. The proposed method is focusing on interpretation of sign language in proper English sentence. Different NLP techniques are used in addition to sign recognition. Input is given as video of sign language followed by framing & segmentation on video. CamShift algorithm is used for tracking & P2DHMM for hand tracking. Haar Cascade classifier is used for sign identification. After sign recognition, the continuous words for respective sign are given as input to POS tagging module. Word net POS tagger is used which is having its own WordNet dictionary. At last LALR parser is used to frame the sentence. In this way proposed sign language interpreter model gives the output in meaningful English sentence & with a goodaccuracy.

***Keywords***: *Sign language, hearing disable, visual interaction, image processing, POS, parse.*

***Mathematics Subject Classification 2010:*** *68M12*

## 1. Introduction

The way of transferring the information from one person to another is called communication. Most of the time people use signs and words for the communication. Natural language is used by normal people to communicate/interact with each other while tactile sign language is used by deaf and dumb people to interact [4]. Nowadays people with disabilities experience difficulties to stand in the race because of ferocious competition in every field. The effort is to develop an application which

*E-mail: sampadawazalwar@gmail.com

will help deaf and dumb people interact with a normal person. According to a survey, India consist of nearly 2.4 million deaf, dumb populations which approximately make up 20% of the world's total deaf and dumb population[20]. For hassle-free interaction between the normal person and deaf and dumb person, there is a need of an interpreter (Person who has the knowledge of sign language, as well as normal language). Sign language is divided into two i.e. Visual Sign Language & Tactile Sign Language.

a) *Visual sign language:* It is used by hearing & speech impaired people
b) *Tactile sign language:* It is used by hearing & sight impaired people.

We are basically working on the visual sign language used by deaf & dumb. Sign Language varies country to country it depends on its culture as Sign language in India is ISL (Indian Sign Language), America uses ASL (American Sign Language), China uses CSL (Chinese Sign Language).

Sign Language is a method of communication for deaf & dumb which is composed of various gestures formed by hand shapes, body orientation & facial expression. Each gesture has a meaning assigned to it. Alphabets in sign language are composed of different hand shapes & words are composed of hand shapes with orientation. Complete visual sign language also includes facial expressions. Visual sign language is an effective means of communication for deaf & dumb [10]. Though it is true, the hearing-impaired have to challenge communication obstacles in a mostly hearing capable society [6]. This research work will concentrate on Visual Sign language interaction.

Natural language is a skill used for understanding human language [10]. It is a part of linguistics and Artificial Intelligence. NLP (Natural language processing) is a step for developing a system that can convert the text (words) in human language. POS tagging is the method of NLP and first introduced in 1960. It is an important method for language processing. For many NLP applications it is the simplest and most stable step. Part Of Speech tagging is the initial step for machine translation, retrieval of information and etc [15]. Second important method in NLP is parsing. Parsing is the method which is followed by the compiler. When we thought of sign recognition, we should consider the major challenges in sign recognition which are mentioned below.

**Major Challenges in Sign Recognition**

Sign Language recognition is not an easy task, looking towards the survey there are so many difficulties. Static sign recognition is comparatively easy than continuous sign recognition. Continuous recognition includes the following problems

a) Identify the start & end of a single sign
b) Sign making speed also varies person to person.
c) The number signs in a phrase vary continuously.
d) Background of the signer & motion of signer also affects the recognition system
e) English language is ambiguous thus it may cause the misinterpretation.
f) Some person has more or less body structure & his emotional quality.

When we think of making a sign language recognition system above all problems should be taken into consideration.

The aim is to develop a Sign Language interaction translator that eases the communication of the deaf people with normal ones. This translator will capture a video of deaf person while interaction and will identify enacted words. These words are then used to frame sentence using Natural Language Processing techniques. Today's research mainly focuses on semi-supervised Machine Learning algorithms.

## 2. Implementation

Our aim is to develop a system which interprets the sign language in English sentences. All existing systems [2, 3, 4, 5, 8] only focuses on recognizing the words which can be interpreted in wrong sentences so the ultimate goal is to convert the recognized continuous sign into proper English sentences. Proposed system details are as follows

The system Model is mainly divided into two phases as follows:

A. Sign language conversion into text i.e. words.

B. Forming meaningful sentence of text using NLP techniques.

**A. PHASE I:** Sign language conversion into text i.e. words.

Moreover work has been done in Sign language detection and conversion. This was done by using image processing techniques in combination with machine learning approach. In proposed work the sign language is detected and converted using following image processing techniques:

1. Input video
2. Framing (Key frame extraction)
3. Segmentation
4. Tracking



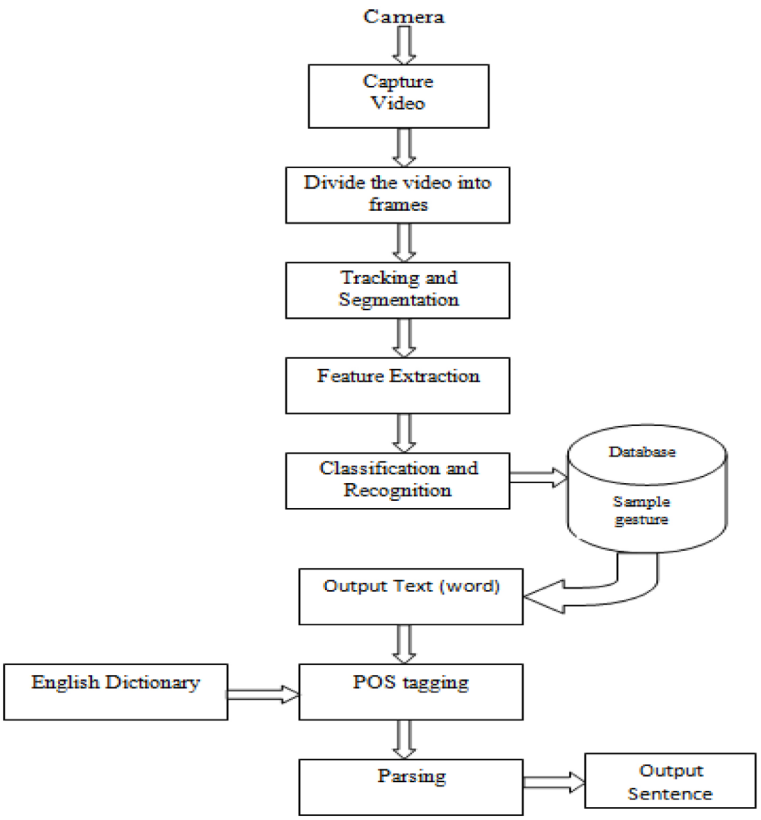



**Fig .1**

**System Flowchart**

5. Feature extraction
6. Classification and Recognition

1. Input video

In this 2D camera upto 12 MP is used to capture the video. It provides the depth as well as skeleton information. 2D camera also reduces the cost of the process. This work does not address hand overlapping issue. The video taken by camera is further divided into frames. There are different type of video formats such as .avi, .mov, .wmv, .mng etc. We have done the work on .avi format video with maximum size of 8 MB.

*Why .avi?*

AVI is Audio Video Interleaved. We are using .avi video format because it supports any audio & video coding formats. It uses RIFF i.e. Resource Interchange File Format (RIFF) which is a generic file container format that is used for storing data in tagged chunks.

2. Framing

Framing of video i.e. given video is divided into frames using formula: Number of frames= Actual video length *FPS

Where, FPS is frame per second rate which is varying from video to video. According to formula elapsed minutes is converted into total seconds and then multiply that by FPS (Frame per Second rate) either by 24Fps, 25Fps, or 30Fps (video in USA). Or you can make FPS to Decimal breakdown of frames. Here in proposed work 1 to 3 sec of videos are taken and the Frame per Second rate is also vary from frame to frame such as 15FPS to 24 Fps. Video of 5 sec with 12 FPS is divided into approx. 56 frames.

3. Segmentation

Various approaches and techniques for segmentation are available in literature [17] but results changes from place to place because various parameters like limited light condition, skin color, background of image etc. Pixel and region based techniques are available [18]. Due to large variation in skin color of Indian people, finding out an adaptive color model for color detection is big challenge. To avoid this limitation HSV (Hue Saturation value) model for segmentation is used. It converts RGB to HSV value using following formula:

$$R' = R/255,\ G'=G/255,\ B'=B/255$$

$$Cmax = \max(R',G',B'),\ Cmin=\min(R',G',B'),\ d=Cmax-Cmin$$

$$H = 60*(G'-B'/d\ *\mathrm{mod}6),\ S=d/Cmax\ ,\ V=Cmax$$

4. Tracking

Tracking algorithm [21] gives information of the position of hand such as color tracking, motion tracking, template matching, blob tracking. Multiple cues integrating methods are available. In proposed work Camshift algorithm is used in which it finds the mean position and track the hands gestures of the frame. Algorithm for tracking is as follows:

CAMSHIFT Algorithm [35]:

1. Choose the position of initial search window i.e. *W* for size *s*.
2. The initial search window is focus at data point $p_k$ and store zeroth moment.
3. Compute the mean location in the search window (*W*)

$$p_k(w) = \frac{1}{|w|}\sum_{j\in w} p_j \tag{1}$$

The mean shift climbs to the gradient

$$p_k(w) - p_k = \frac{f'(p_k)}{f(p_k)} \tag{2}$$

Focus the search window on mean location computed in step3

4. Until convergence repeat 3 and 4.

Near the mode $f'(p) = 0$ so the mean shift algorithm converges there. Zeroth moment

$$M_{\infty} = \Sigma_x \Sigma_y I(x,y) \tag{3}$$

Where, $I(x, y)$ is intensity (pixel value) of position $x$ and $y$ in the image.

5. Feature Extraction

In order to implement a continuous forward search method and sequential composition of gesture models, the first mentioned type has been used in this research [30]. There are three types of parameters in the P2DHMMs (Pseudo-2D hidden Markov models). Since the hand image is 2D, we further divided the Markov

transition parameters into state transition probabilities and super-state transition [9]; each is denoted as

$$\overline{\alpha_{kl}} = p(r_{t+1} = l \mid r_t = k),\ 1 <= k, 1 <= N \text{ and}$$

$$\alpha_{ij} = p(q_{t+1} = j \mid q_t = i),\ 1 <= I, k <= M$$

Where, $r_t$ denotes a super-state which coincide with a HMMs $\lambda_k$ and $q_t$ denotes a state observing at time $t$. The mode has N super-states and the HMMs $\lambda k$ [30], is defined as standard HMM consisting of $M$ states.

**Pseudo 2 Dimensional Hidden Markov Model Algorithm:**

The DCT coefficient of the images is used followed by the selection of images with the most distinguishable coefficients for P2-DHMMs training model. So that from the obtained data set, one can extracts the most important information available in subject's images [32]. The algorithm used for determining the best training images is as follows [7]:

1. Select a training image arbitrarily
2. Calculate the distance between the DCT vector of other images and the DCT vector of the chosen image;
3. Select as the second training image, the image that has the biggest distance with the first chosen training image;
4. For all images, obtain the overall distance between each image and selected training images using equation;

$$D_{i,j}^2 = \sum_{n=1}^{N} (d_i(n) - d_j(n))^2 \tag{4}$$

   Index of the next training image = $argmax_i$ (min ($D_{i,j}$)) where '$N$' is the length of image vector (No. of rows × No. of columns).
5. Choose another training image, the image with the biggest distance;
6. If there is still training image, go to 4.
7. End.

## 6. Classification and Recognition

Present work concentrates on Haar Cascade Classifier Algorithm which is used for getting the region of interest from hand image preceding preprocessing techniques involving skin detection and size normalization.

**Training stage:**

Haar Cascade Classifier is trained using multiple image samples of each gesture [35]. These samples are stored in the database. These samples are collected from different people with different hand color, shape and different lightening condition in different background. Accuracy of recognition is done by locating area of interest i.e. hands of person in each sample images. This can be done by drawing a box around the hand feature. The co-ordinates of interested region in samples are then analyzed to measure the contrast between each of these images. This stage will use to build the required cascade and find thresholds after analyzing the coordinates of hand sign. Classifier uses Haar like feature using simple Haar function given in formula as follows [36]:

$$H(t) = 1 \; 0 \le t \le ½ \qquad (5)$$
$$= -1 \; 1/2 \le t \le 1$$
$$= 0 \text{ otherwise}$$

**Testing Stage:**

After training stage, classifier is now trained to distinguish between different signs. Testing is performed on the video taken through 2D camera. Output of this phase is in word format. The output Words are I and Active because deaf people didn't use the verbs like am, is, are etc. Let us denote the output words as *M*.

B. PHASE II: Forming meaningful sentence of text using NLP techniques

This phase is concerned with meaningful sentence formation using group of words obtain from above image processing techniques. The sentences are formed using various POS tagging, Parsing and grammar.

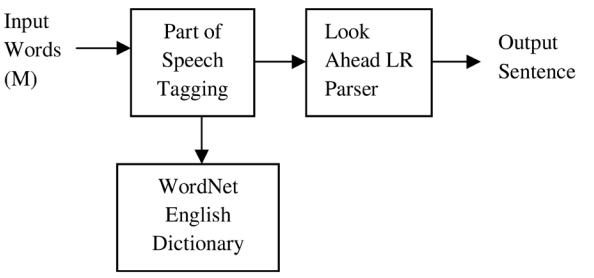



**Fig. 2**

**NLP model for framing sentences**

1. POS Tagging:

The words obtain from phase I passed to NLP The tag may indicate one of the parts-of-speech tag like noun (N), preposition (Prep), pronoun, adjective (Adj), verb (V), adverb (Adv), conjunction, and interjection [29]. The input is tokenized and a training dataset is used for detecting the equivalent part of speech of each word in the sentence.

WordNet POS tagging Algorithm:

1. Take *M* as input and specified list of POS tags such as N,V, Adj, Adv, Prep.
2. Search the words for specified tags in the training corpus.
3. Find the probability of tags for words in given corpus as follows:

$$P(ti \mid w) = c(w, ti)/(c(w, t1) + \ldots + c(w, tk))$$

Where $w$ = input word $M$, $ti$ = tag for input word $M$, $c(w,ti)$ is number of times appear in corpus

4. Tag the word with maximum probability.

2. Designing of grammar

Grammar specifies the set of rules to define legal sentences in particular language. These rules work together for constructing the valid sentences. Simple and easiest form of grammar is context free grammar [13]. It is easy to deal and write the context free grammar for people. Most of the time grammar is compromised by writing rules in context free style. These rules are then augmented to deal with some kind of complications. For grammar construction variety of approaches are used. It is useful to examine the use in following ways:

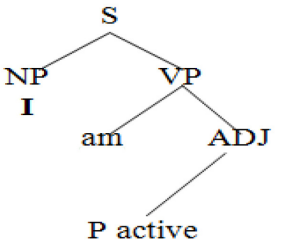
1. Conceptual tool- It is used for finding complexity of natural language by capturing and reporting.
2. Formal notation – It is used to trace syntactical and semantical feature of language which is used by parser.

Representation of Grammar design is as follows [34]:G = {$V$, $\Sigma$, $P$, $S$}

3. Parsing Technique:

Bottom-up parsing is previously known as parsing algorithm and it is used in the shift-reduce parsers which is common for computer languages [16] .In bottom-up parsing, the parser starts with the input words, and tries to build trees from the words in upward direction, by applying rules

from the grammar one at a time [27]. The parse is successful if the parser succeeds in structure a tree which is rooted in the start symbol (S) and covers all of the input [25]. Figure show the bottom-up search space, beginning with the sentence I Active. The parser begins by looking up each word in the lexicon and build tree with the part of speech for each word.LALR parser is simplified version of LR parser it is used to separate and analyze the text according to their production rules specified by formal grammar. The input to the parser is pronoun, verb, conjunction, adjectives etc. Tactile sign language does not consists words like is, am, are and article. So to construct meaningful sentence parse tree is used which insert word and articles to the appropriate places. Here in this project Standford tagger with combination of LALR parser for forming a meaningful sentence are used. Bottom Up LR parser is shown as follows:

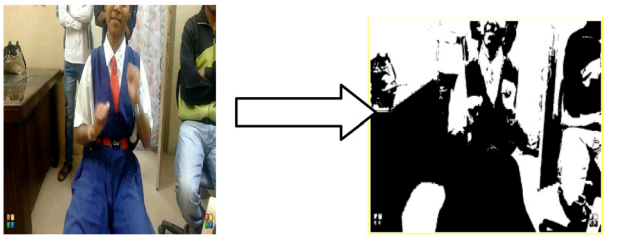



## 3. Result Discussion and Analysis

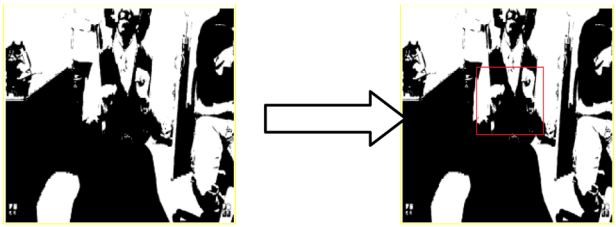
**Step 1:** Take video as an input. Specifications of video is as follows: Type of file: .avi, Size: 7.49mb, Length: 5 sec, Frame width: 900, Frame height: 508, Frame rate: 12 Fps

**Step2:** Divide the video into frames using formula: Number of frames= frame time*FPS

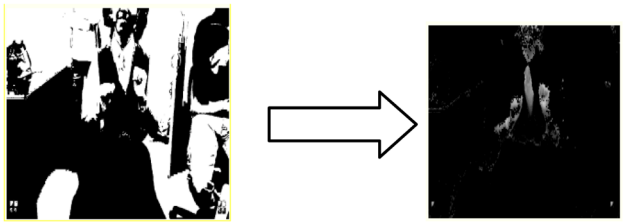
**Step 3:** Binaries the frames to convert it into Grayscale.



**Step 4 :** Segmentation and Tracking performed done using HSV and CAMSHIFT algorithm.



**Step 5:**Feature extraction is done using P2DHMM algorithm which tracks the feature of the given frame.



**Step 6:** Classification and recognition of hand movement is completed using Haar Classifier which identify the Feature using Database which consist 33 samples of model feature expression and output word is obtained. The outputs words are "I" and "Active". Sample of words used are : I, sorry, engineer, afternoon, clean, floor. How. You, happy, sad, enjoy, father, mother etc. The samples are taken from the hearing impaired person school of Shankar nagar Nagpur under observation of the experts.

**Step 7:** The output words is then given to NLP engine consist of POS tagger and LALR parser. POS tagger tags the words and Parser is use for sentence formation of the outputted words. POS tagging and parsing of above words is as follows: I(N), Active(Adj)

**Analysis:**

Above table shows the list of some of videos which were given as an input to the system. All these videos are of different duration. Total 60 videos are taken with different background and in different light conditions. No. of frames varies with the size of video The accuracy of results of sign language conversion system depends on the database sample, video resolution.

**Table 1**
**Analysis of frame recognition**

| Video name | Input video sentence | Length of video | No. of Frames | Correctly Recognized frames |
|---|---|---|---|---|
| V1.avi | Clean Floor | 3 sec | 42 | 38 |
| V2.avi | I am active | 5 sec | 60 | 40 |
| V3.avi | I am Engineer | 4 sec | 50 | 46 |
| V4.avi | I am sorry | 2 sec | 24 | 24 |
| V5.avi | I love my mother | 5 sec | 60 | 42 |

**Table 2**
**Comparison of System Generated Output with Manually Generated Output**

| System generated output | Manually generated output |
|---|---|
| I am sorry | Sorry |
| Clean Floor | Clean floor |
| I am Engineer | Engineer |
| Please be strong | Please Strong |
| You are welcome | You Welcome |



**Subjective Analysis**

For analysis purpose the sentence obtained from videos checked with the deaf and dumb expert. Output of Sign language conversion system for ten videos with different length gives five sentences which was examined by school teacher of deaf dumb school who interpret the same five video manually. According to that the output of this system and manual interpretation are very similar. So the system gives nearly 90% of accuracy.

## 4. Conclusion and Future Scope

The major objective of this paper is to give significance of sign language interpretation and focus on conversion of deaf/dumb sign

language into sentences to ease their day to day life. This work was done on a set of limited database words & tested for small sentences. In future we are trying to extend our work by considering the large database which will cover maximum words & try to develop algorithms in NLP which will work on long sentences. In future to improve the efficiency of sign recognition, facial expression detection can be included in sign recognition module, which will help us to punctuate the sentence correctly. Which will intern help us to identify the emotion of person. If we follow some linguistic principles and go forward with the help of Sign Language trainer and deaf people, this work can really give justice for implementation of sign language interpretation system. The assistive system for deaf can link the communication breach between hearing disable and normal people without isolating them in the society.



**References**

[1] Chetan Arora, Mehrdad Sabetzadeh, Lionel Briand and Frank Zimmer, "Automated Checking of Conformance to Requirements Templates using Natural Language Processing", *IEEE transactions on Software Engineering*, Volume: ,PP.1-25, MAY 2015.

[2] Shreyashi Narayan Sawant,M. S. Kumbhar, "Real Time Sign Language Recognition using PCA", *2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*.

[3] Geetha.M, Aswathi.P.V, Amrita VishwaVidyapeetham, "Dynamic Gesture recognition of Indian Sign Language considering Local motion of hand using Spatial location of Key Maximum Curvature Points", *2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*.

[4] Maryam Pahlevanzadeh, Mansour Vafadoost, Majid Shahnazi, "Sign Language Recognition" , Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013).

[5] Ziqi Wang, Gu Xu, Hang Li,Ming Zhang, " A Probabilistic Approach to String Transformation", *IEEE Transactions On Knowledge And Data Engineering*, VOL. 26, NO. 5 1063 - 1075, MAY 2014.

[6] Kanchan Dabre,Surekha Dholay, "Machine Learning Model for Sign Language Interpretation using Webcam Images", *2014 IEEE International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA)*.

[7] A. Wilson and A. Bobick, "Parametric Hidden Markov Models for Gesture Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 9, pp. 884-900, Sept. 1999.

[8] Jonathan Alon, Quan Yuan, Stan Sclaroff, "A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation" *IEEE Transactions on pattern analysis and machine intelligence*, vol. 31, no. 9, september 2009.

[9] S. Tsagaris, S. Manitsari, "Color Spaces Comparisons for Skin Detection in Finger Gesture Recognition", *International Journal of Advances in Engineering & Technology,* 2013, pp. 1431-1441.

[10] Mohamed Mohandes, Junzhao Liu, Mohamed Deriche, "A Survey of Image-Based Arabic Sign Language Recognition", 2014 11th International on Multi-conference on systems, signals and devices (SSD).

[11] Chandhana Surabhi.M, "Natural Language Processing Future", Proceedings of International Conference on Optical Imaging Sensor and Security, Coimbatore, Tamil Nadu, India, July 2-3, 2013.

[12] Divya S, Kiruthika ,S Nivin Anton A L and Padmavathi S, "Segmentation, Tracking And Feature Extraction For Indian Sign Language Recognition", *International Journal on Computational Sciences & Applications (IJCSA)* Vol.4, No.2, April 2014. Listening deaf through Tactile sign language

[13] Suvarna G Kanakaraddi,V Ramaswamy, "Natural Language Parsing using Fuzzy Simple LR (FSLR) Parser", *2014 IEEE Advance Computing Conference(IACC).*

[14] Matthias Zimmermann, Jean-Cedric Chappelier,Horst Bunke, "Offline Grammar-Based Recognition of Handwritten Sentences", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* VOL. 28, NO. 5, May 2006.

[15] Eduardo Blanco and Dan Moldovan, A Semantic Logic-Based Approach to Determine Textual Similarity, IEEE/ACM Transactions On Audio, Speech, And Language Processing, VOL. 23, NO. 4,683-693, APRIL 2015.

[16] Roberto Navigli and Paola Velardi,"Structural Semantic Interconnections: A Knowledge-Based Approach to Word Sense Disambiguation", *IEEE transactions on Pattern Analysis And Machine Intelligence*, VOL. 27, NO. 7, JULY 2005.

[17] L. Howe, F. Wong, A. Chekima, "Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition", *Information Technology, ITSIM,* IEEE-978-4244-2328-6, 2008, pp.1-7.

[18] S. Phung, A. Bouzerdoum and D. Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison", *IEEE Transaction on Pattern Analysis an Machine Intelligence*, Vol.27, No. 1, 2005, pp. 148-154.

[19] C. Jung, C. Kim, S. Chae, and S. Oh, "Unsupervised Segmentation of Overlapped Nuclei Using Bayesian Classification", *IEEE Transaction on Biomedical Engineering,* Vol. 57, No.12,Dec-2010, pp. 2825-2832

[20] A. Chitade, S. Katiyar, "Color Based Image Segmentation Using K-means Clustering", *International Journal of Engineering Science and Technology,* Vol.2, No.10, 2010, pp. 5319-5325.

[21] A. Yilmaz, O. Javed, M. Shah, "Object Tracking: A Survey", ACM Computing Surveys,Vol. 38, No. 4, Article 13, December 2006, pp. 1-45

[22] A. S. Ghotkar, G. K. Kharate, "Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction", *International Journal of Human Computer Interaction (IJHCI), Computer Science Journal,* Malaysia, Volume 3, no. 1, ISSN 2180-1347, April 2012, pp. 15-25.

[23] L. Yun, Z. Peng, "An Automatic Hand Gesture Recognition System based on viola-Jones Method and SVMS", International workshop on Computer Science and Engineering, *IEEE Computer Society*, 2009, pp. 72-76.

[24] D. Zhang, C. Lu, "Review of shape representation and description techniques", *The Journal of the Pattern Recognition Society*, Elsevier, 2004, pp. 1-19.

[25] A. K. Jain, R. Duin, Mao, "Statistical Pattern Recognition: A Review", IEEE Transactions On Pattern Analysis And Machine Intelligence, Volume 22, No. 1, January 2000. pp. 4-37.

[26] J. Li, B. Lu, "An adaptive image Euclidean distance", *Pattern Recognition Journal,* Elsevier,Volume 42, 2009, pp. 349 -357.

[27] Guo-Dong, A. K. Jain, W. Ma, H. Zhang, "Learning similarity Measure for Natural Image Retrieval with Relevance Feedback", *IEEE Transactions on Neural Networks,* Vol.13, No. 4, July2002, pp. 811-820.

[28] K. K. Wong, R. Cipolla, "Continuous gesture recognition using a sparse Bayesian classifier", *International conference on pattern recognition*, 2006, pp. 1084-1087.

[29] N. D. Binh, E. Shuichi, T. Ejima, "Real time Hand Tracking and Gesture Recognition System", ICGST International Conference on Graphics, Vision and Image Processing, GVIP 05 Conference, Egypt, Dec-2005, pp. 362-368..

[30] A. Corradini, "Real-Time Gesture Recognition by means of Hybrid Recognizers", GW 2001,LNAI 2298, Springer-Verlag Berlin Heidelberg 2002, pp. 34-47.

[31] P. Bao, N. Binh, T. Khoa, "A new Approach To Hand Tracking and Gesture Recognition by a New Feature Type And HMM", International Conference on Fuzzy Systems and Knowledge Discovery, IEEE Computer Society, 2009, pp. 3-6.

[32] Guo-Dong, A. K. Jain, W. Ma, H. Zhang, "Learning similarity Measure for Natural Image Retrieval with Relevance Feedback", *IEEE Transactions on Neural Networks*,Vol.13, No.4, July 2002, pp. 811-820.

[33] P. Garg, N. Agrawal, S. Sofat, "Vision based Hand Gesture Recognition", Proceedings of world Academy of Science, Engineering and Technology, Vol.37, 2009, pp. 1024-1029.

[34] Sumeet R. Agarwal, Sagarkumar B. Agrawal, Akhtar M. Latif, "Sentence Formation in NLP Engine on the Basis of Indian Sign Language using Hand Gestures", *International Journal of Computer Applications* (0975 – 8887) Volume 116 – No. 17, April 2015.

[35] Miss Sulochana M Nadegeri,Dr.S D Sawarkar, Mr.A.D Gawande, "Hand Gesture Recognition Using CAMSHIFT Algorithm, Third International Conference on Emerging Trends in Engineering and Technology.

[36] Ogwu, E. J., M. Talib, and O. A. Odejobi. "Text-to-speech processing using African language as case study." *Journal of Discrete Mathematical Sciences and Cryptography* 9.2, 2006, pp. 365-382.