

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

## Transforming Indian Sign Language into Text Using Leap Motion

P.Karthick<sup>1</sup>, N.Prathiba<sup>2</sup>, V.B.Rekha<sup>3</sup>, S.Thanalaxmi<sup>4</sup>

Assistant Professor, Department of Computer Engineering, Anand Institute of Higher Technology, Chennai, India<sup>1</sup>

U.G. Student, Department of Computer Engineering, Anand Institute of Higher Technology, Chennai, India<sup>2</sup>

U.G. Student, Department of Computer Engineering, Anand Institute of Higher Technology, Chennai, India<sup>3</sup>

U.G. Student, Department of Computer Engineering, Anand Institute of Higher Technology, Chennai, India<sup>4</sup>

**Abstract:** A sign language looks up the manual communication and body language to convey meaning, as opposed to acoustically conveyed sound patterns, which involve simultaneous combination of hand shapes, orientation and movement of hands, arms or body, and facial expressions to fluidly express a speaker's thought. The Leap device tracks the data like point, wave, reach, grab which is generated by a leap motion controller. The system implements DTW combined with IS algorithm for converting the hand gestures into an appropriate text, aided by leap device that consists of inbuilt camera and two IR sensor to capture hand signals. Neuro Linguistic Programming (NLP), a division of Artificial Intelligence which includes Natural Language Processing and Neural Networks. The IS invokes a trigger when the current environment changes dynamically, DTW handles gesture transformation mapped with similar patterns.

**Keywords:** Neuro Linguistic Programming (NLP), Dynamic Time Warping (DTW), Intelligence Sense (IS), Infrared Rays (IR)

### I. INTRODUCTION

Artificial Intelligence is the branch of computer science concerned with making computers behave like humans. The term was coined in 1956 by John McCarthy at the Massachusetts Institute of Technology. Currently, no computers exhibit full artificial intelligence. The greatest advances have occurred in the field of games playing. In the area of robotics, computers are now widely used in assembly plants, but they are capable only of very limited tasks. Robots have great difficulty identifying objects based on appearance or feel, and they still move and handle objects clumsily. Natural-language processing offers the greatest potential rewards because it would allow people to interact with computers without needing any specialized knowledge. You could simply walk up to a computer and talk to it. Unfortunately, programming computers to understand natural languages has proved to be more difficult than originally thought. Some rudimentary translation systems that translate from one human language to another are in existence, but they are not nearly as good as human translators. There are also voice recognition systems that can convert spoken sounds into written words, but they do not understand what they are writing; they simply take dictation. Even these systems are quite limited -- you must speak slowly and distinctly. Today, the hottest area of artificial intelligence is neural networks, which are proving successful in a number of disciplines such as voice recognition and natural language processing. Artificial neural networks are computational models inspired by animal central nervous systems (in particular the brain) that are capable of machine learning and pattern recognition. They are usually presented as systems of interconnected "neurons" that can compute values from inputs by feeding information through the network. Like other machine learning methods, neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming, including computer vision and speech recognition. Warren McCulloch and Walter Pitts created a computational model for neural networks based on mathematics and algorithms. They called this model threshold logic. The model paved the way for neural network research to split into two distinct approaches. One approach focused on biological processes in the brain and the other focused on the application of neural networks to artificial intelligence. The proposed system understands the human gestures and converts them in to meaningful sentences using Leap Motion Controller. The Leap Motion controller is a small USB peripheral device which is

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

designed to be placed on a physical desktop, facing upward developed by Leap Motion Inc. Leap Motion, Inc. is an American company that manufactures and markets a computer hardware sensor device that supports hand and finger motions as input, analogous to a mouse, but requiring no hand contact or touching. Using two monochromatic IR cameras and three infrared LEDs, the device observes a roughly hemispherical area, to a distance of about 1 meter (3 feet). The LEDs generate a 3D pattern of dots of IR light and the cameras generate almost 300 frames per second of reflected data, which is then sent through a USB cable to the host computer, where it is analysed by the Leap Motion controller software using "complex math" in a way that has not been disclosed by the company, in some way synthesizing 3D position data by comparing the 2D frames generated by the two cameras. Leap tracks the hand as a general object, allowing for actions like pinching, crossing fingers, moving one hand over another and hand-to-hand interactions like brushing and tapping fingers on one another. The API combines the use of Neural Networks to track moving data efficiently. This project makes a combined effort of involving all new technologies including NLP, AI, and Neural Networks along with the Leap Motion API for successful completion. The Leap analyses the overall motion which occurred since an earlier frame and synthesizes representative translation, rotation, and scale factors. For example, if you move both hands to the left in the Leap field of view the frame contains translation. If you twist your hands as if turning a ball, the frame contains rotation. If you move your hands towards or away from each other, the frame contains scaling. If it only detects one hand, then the Leap bases the frame motion factors on the movement of that hand. If it detects two hands, then the Leap bases the frame motion factors on the movement of both hands together. You can also get independent motion factors for each hand from a Hand object. Frame motions are derived by comparing the current frame with a specified earlier frame. The attributes describing the synthesized motion include: You can apply the motion factors to manipulate objects in your application's scene without having to track individual hands and fingers over multiple frames. The Leap API provides as much information about a hand as possible. However, the Leap may not be able to determine all hand attributes in every frame. For example, when a hand is clenched into a fist, its fingers are not visible to the Leap so the finger list will be empty. Your code should handle the cases where an attribute in the model is not available. The Leap recognizes certain movement patterns as gestures which could indicate a user intent or command. The Leap reports gestures observed in a frame the in the same way that it reports other motion tracking data like fingers and hands. For each gesture observed, the Leap adds a Gesture object to the frame. You can get these Gesture objects from the Frame gestures list. The Leap Motion Controller senses how you naturally move your hands and let you use your computer in a whole new way. Point, wave, reach, grab. Pick something up and move it. Do things you never dreamed possible.



Fig 1: Leap device

## II. RELATED WORKS

A Sign Language Recognition (SLR) system has to be designed to recognize a hand gesture. Gestures in sign language are defined as specific patterns or movements of the hands, face or body to make out expressions. The sign capturing methods are vision based sign extraction, data glove and Electromyography. The continuous sign recognition is very complex as it cannot be separated as the speech is separated. Hence continuous sign recognition system uses Hidden Markov Model (HMM) and Electromyography (EMG) segmentation [1]. Sign Language, the natural communication

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

medium for a deaf person, is difficult to learn for the general population. The prospective signer should learn specific hand gestures in coordination with head motion, facial expression and body posture. Sign tutor helps the user by providing them with sign videos, text-based description, pictures of hand gestures and 3D animated avatar [2]. A vision-based static hand gesture recognition algorithm consists of three stages: pre-processing, feature extraction and classification. The pre-processing stage involves following three sub-stages: segmentation which segments hand region from its background images using a histogram based thresholding algorithm and transforms into binary silhouette; rotation that rotates segmented gesture to make the algorithm, rotation invariant; filtering that effectively removes background noise and object noise from binary image by morphological filtering technique. A localized contour sequence (LCS) based feature is used here to classify the hand gestures. A k-mean based radial basis function neural network (RBFNN) is also proposed here for classification of hand gestures from LCS based feature set [3]. Microsoft Kinect sensor plays a vital role in the sensing and robotics communities. However, the applications are programmed are generated in C alone. To overcome this issues, the VU- Kinect Simulink, an integration of kinect and simulink has been developed which provides easy access to the sensor’s camera and depth image streams. The VU- Kinect block is used to track a 3-D object [4]. Vision based recognition system consist of three main components: hand gesture modelling, hand gesture analysis and hand gesture recognition. Gesture model describes how the hand gesture is to be represented. Analysis is performed to compute the model parameter from the image feature. The analysis phase is followed by recognition phase that classify the model prepared. Hand gesture recognition consists of feature detection, hand localization, feature extraction and parameter computation [5]. Classification technique deals with the Euclidean distance metric. Translated ISL is displayed with the help of a 3D virtual human avatar. Input to the system is the clerk’s speech which is in English. The speech recognition module recognizes speech and makes a text output. The output from the parser is given to an eliminator module which performs a reduction task by eliminating unwanted elements and further the root form of verbs are found using the stemmer module. The structural divergence of English and ISL is handled by a phrase reordering module using ISL dictionary and rules [6].

### III. OVERALL ARCHITECTURE

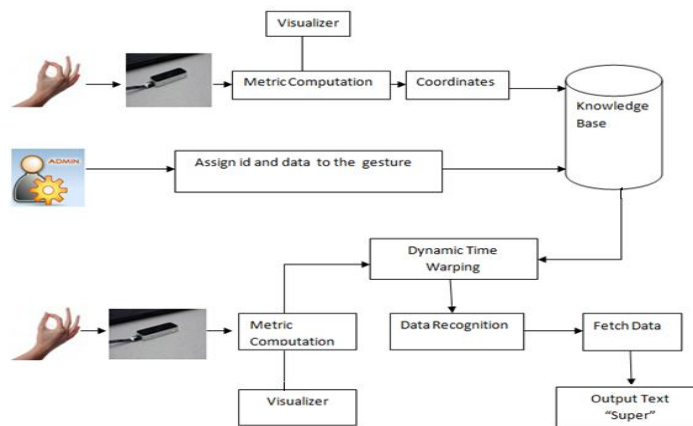


Fig. 2 Overall architecture

### IV. MOTION TRACKING

The Leap software analyses the objects observed in the device field of view. It recognizes hands, fingers, and tools, reporting discrete positions, gestures, and motion. Gestures are recognized through a tool called Leap Visualizer. It is used to visualize the Gesture recognition to view motion tracking data generated by leap motion controller. As the Leap tracks hands, fingers, and tools in its field of view, it provides updates as a set, or frame, of data. Each frame contains lists of the basic tracking data, such as hands, fingers, and tools, as well as recognized gestures and factors describing the overall motion in the scene.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

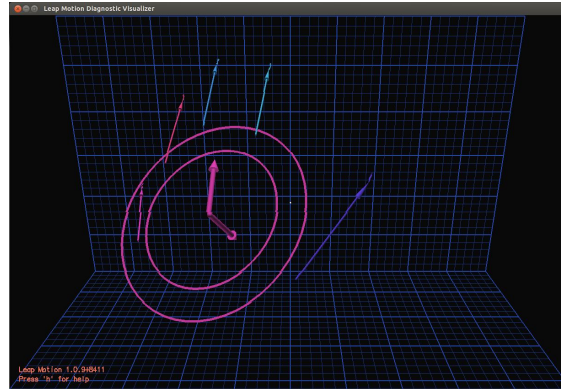


Fig. 3 Leap Visualizer

## V. GESTURE LEARNING

The gesture recognized through visualizer will be captured along with text, computed with DTW algorithm, and added to KB, where KB is the knowledge base that contains metric computation of each gesture. When it detects a hand, finger, tool, or gesture, the Leap assigns it a unique ID designator. The ID remains the same as long as that entity remains visible within the device's field of view. If tracking is lost and regained, the Leap may assign a new ID (the software may not know that the hand or finger is the same as the one visible earlier).

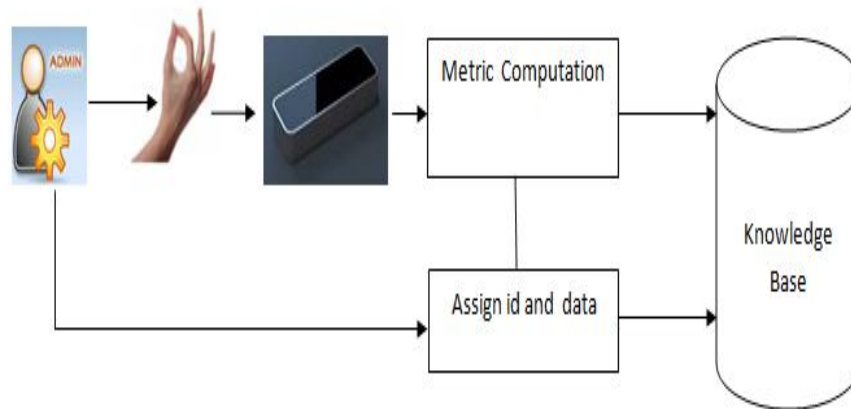


Fig. 4 Gesture Learning

## VI. GESTURE RECOGNITION

The gestures were given as input and the corresponding text assigned to that gesture will be retrieved from the knowledge base and db. The system implements DTW combined with IS algorithm for converting the hand gestures into an appropriate text, aided by leap device. The DTW algorithm finds an optimal match between 2 time series. The 2 time series' data are nonlinearly warped in such a way that the similar regions are aligned and a minimum distance

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

between them is obtained. DTW works by warping the time axis iteratively until an optimal match between the 2 sequences is found.

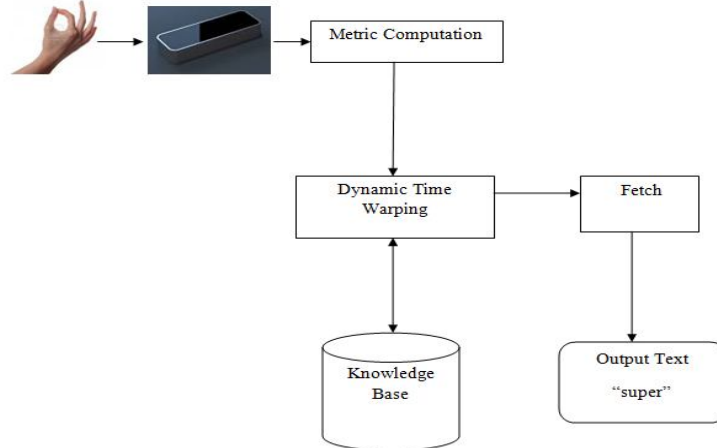


Fig. 5 Gesture Recognition

## VI. CONCLUSION

The system has been trained to find out the textual context for the user provided gestures with less noise. Efficient knowledgebase has been created with hand gestures and matching text/description using competent DTW algorithm, fetch the textual context from knowledgebase for the user's gestures. The System discards an inappropriate gesture that does not match with the knowledgebase.

## REFERENCES

- [1] Al-Ahdal, M and Tahir, N. (2012) 'Review in sign language recognition systems', IEEE Symposium on Computers and Informatics (ISCI), pp. 52-57.
- [2] Aran, O and Ari, S.(2009) 'SignTutor: An interactive system for sign language tutoring', IEEE Multimedia, vol.16, issue 1, pp. 81 – 93.
- [3] Ghosh, D and Ari, S. (2011) 'A static hand gesture recognition algorithm using K-Mean based Radial Basis Function Neural Networ', 8<sup>th</sup> International Conference on Information, Communications and Signal Processing (ICICS11), pp. 1 – 5.
- [4] Joshua Fabian, Tyler Young, James, Peyton Jones and Garrett M. Clayton, (2014) 'Integrating the Microsoft Kinect with Simulink: Real-Time ObjectTracking Example', Mechatronics, IEEE/ASME Transactions on (Volume:19 , Issue: 1 )".
- [5] Kang, S, Nam, M and Rhee,P. (2008) 'Colour based hand and finger detection technology for user interaction', International Conference on Convergence and Hybrid Information Technology, pp. 229 – 236.
- [6] Rekha, J, Bhattacharya, J and Majumder, S. (2011) 'Shape, texture and local movement hand gesture features for Indian Sign Language recognition', 3rd International Conference on Trendz in Information Sciences and Computing (TISC), pp. 30 – 35.
- [7] Kang, S, Nam, M and Rhee,P. (2008) 'Colour based hand and finger detection technology for user interaction', International Conference on Convergence and Hybrid Information Technology, pp. 229 – 236.
- [8] Lilha, H and Shivmurthy,D. (2011) 'Analysis of pixel level features in recognition of real life dual-handed sign language data set', IEEE International Conference on Recent Trends in Information Systems (ReTIS), pp. 246 – 251.