Dynamic Hand Gesture Recognition for Sign Words and Novel Sentence Interpretation Algorithm for Indian Sign Language Using Microsoft Kinect Sensor

Article · January 2015

DOI: 10.13176/11.626

2 authors:

Archana Ghotkar
Pune Institute Of Computer Technology

Gajanan K. Kharate
Matoshri College of Engineering & Research Centre

Journal of Pattern Recognition Research 1 (2015) 24-38

Received February 20, 2015. Accepted July 10, 2015.

Dynamic Hand Gesture Recognition and Novel Sentence Interpretation Algorithm for Indian Sign Language Using Microsoft Kinect Sensor

Archana S. Ghotkar *archana.ghotkar@gmail.com*
Pune Institute of Computer Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

Gajanan K. Kharate *gkkharate@yahoo.co.in*
Matoshri COE and Research Centre, Savitribai Phule Pune University, Nashik, Maharashtra, India

Abstract

Indian sign language interpretation is an important task to facilitate communication among Indian deaf community and other people. Dynamic hand gesture recognition among other gesture modalities is a major step towards sign language recognition as well as any human computer interaction applications. The main focus of this paper is to design and develop a new algorithm for Indian sign language sentence creation considering limitation of continuous sign language recognition. This paper explores two algorithms for word recognition. Rule based and Dynamic Time Warping-based methods for Indian sign language word recognition are developed. The Dynamic Time Warping-based method gave better accuracy for continuous word recognition than the rule-based method. The proposed new approach for Indian sign language sentence interpretation using inverted indexing overcomes the challenges of conventional continuous sentence recognition in sign language interpretation.

Keywords: Indian sign language, dynamic hand gesture recognition, human computer interaction, dynamic time warping, inverted indexing

1. Introduction

Sign Language (SL) interpretation is a challenging and demanding application among other human computer interaction (HCI) applications. Indian sign language (ISL) is the main mode of communication among deaf Indians. Till 2012, over five million deaf Indians were using ISL [1]. Deaf people, especially those who live in villages, do not have access to sign language. In India, with vast diversity and cultural differences, ISL varies from region to region as communication language. However, in all large towns and cities across the Indian subcontinent, deaf people use sign languages which are not universal sign languages. Extensive work has been done in this regard by creating awareness amongst the sign language teachers in implementing the standardized ISL [2]. Research work on ISL interpretation started late in India due to the lack of standardization. Even now, deaf people are isolated from the society because of communication gap with other people. There are two ways to bring deaf people into the mainstream of society, i) Human interpreter and ii) Computer interpreter. As compared to number of deaf people, human interpreters are much less in number and moreover they may not be available every time. This paper explores use of computer interface to communicate with deaf people. Sign language recognition involves field of computer vision, pattern recognition as well as linguistic study. It is a very challenging problem for the researchers due to the multi-modality gesture interface and limitation of vision based technology. Sign language considers different modality gestures such as lip movements, facial expression, eye brow movements and hand gesture. All ISL alphabet and number signs are composed of hand posture. Approximately 80% of ISL sign words are composed of hand gesture. So, amongst all gesture modalities , hand plays the most important role in ISL interpretation. ISL consists of manual (hand shape, location, orientation and movement) and non-manual (facial expression, lip movement, eye gaze) signs which are further classified into static and dynamic gestures [3]. There are broadly two approaches to deal with gesture recognition, i) hardware based (data glove, data suit) and ii) vision based. Vision based approach is found more suitable and practical as compared to hardware based approach. Though vision based approach is more practical, it is very challenging due to variable lighting condition, dynamic background as well as skin color detection [4]. Vision based processing starts from normal 2D web camera to stereo 3D camera. With the advancement of sensor technology in computer vision field, Microsoft Kinect sensor was used to deal with dynamic hand gesture recognition problem. Various researchers [4] are using Microsoft Kinect sensor for SL recognition system. Its advantages are i) provide depth and skeleton information , ii) ease of use, iii) it is not affected by variation in lighting condition, iv) no constraints of background, and v) quite inexpensive compared to other stereo cameras. In this paper, dynamic hand gesture recognition for ISL word and ISL sentence interpretation is presented. Main focus has been placed on recognizing continuous ISL words which are dynamic in nature and interpreted into text form. Two dynamic hand gesture recognition algorithms were explored and tested on ISL manual dataset for word recognition with single training sample. These recognized text signs were given to sentence interpretation algorithm as an input keyword and possible sentence was interpreted with inverted indexing concept.

The rest of the paper is organized as follows: section 2 covers previous literature work done on dynamic hand gesture recognition for SL interpretation using Kinect sensor. Section 3 presents ISL dataset and grammar model developed for the same. Section 4 presents methodology for dynamic hand gesture recognition for ISL word and sentence. Section 5 presents experimental work for dynamic hand gesture recognition and sentence interpretation. Section 6 describes conclusion drawn and scope of future work.

2. Literature Survey

Various researchers are working on sign language recognition (SLR) such as American sign language [5, 14], Japanese SL [6], Swedish SL [7] and other SLs. Basically, sign language recognition is divided into three parts i) alphabet and number recognition, ii) word recognition and iii) sentence interpretation. In most of the SL, alphabets and numbers are one or two handed static signs while few alphabets require motion. Alphabet recognition mostly requires finger spelling of the proper noun while in some cases it may be part of a word. Words are symbolic signs with combination of manual and non-manual gestures with motion. These symbolic signs are used to convey the message. SL is a highly structural language, with its own morphology, phonology, grammar and is suitable for computer vision algorithms [3]. Sentences in SL are combination of symbolic signs with grammar and sometimes depend mostly on the culture of SL. Therefore, it becomes a challenging task to handle continuous sentence recognition but at the same time it is essential to work on continuous sentence recognition. There are various computer vision and pattern recognition methods available to deal with dynamic hand gesture recognition for SL word and sentence recognition such as,SVM [9, 15], KNN [8],PCNN [11, 16], HMM [5, 10, 17] and DTW [12] and graph matching [11].Most of the researchers were using data glove [5] for continuous SLR. Researchers are now referring RGB-depth camera such as Microsoft Kinect sensor for the SLR over 2D camera because of the image preprocessing involved in the case of latter. In India, research on ISL interpretation started late and very less work is going on at present on ISL continuous word recognition. Kishore and Kumar [13] worked on video based isolated ISL word recognition using fuzzy logic and achieved 96% accuracy. Kalin and Jonas [7] developed educational signing game based on isolated sign recognition of Swedish sign language using Microsoft Kinect. To train the system HMM model was used for a corpus of 51 signs which achieved 89.7% average recognition accuracy. Frank and Sandy [14] used Kinect for interpretation of American sign language for 10 different isolated words. Recognition accuracy of 97% was achieved using support vector machine. Yanhua et al. [6] presented recognition system for Japanese sign language using Microsoft Kinect sensor. A method was developed to employ two Kinects for getting more dataset of hand signs for which point cloud library (PCL) was used to get processed data. Zang et al. [15] used improved SURF algorithm and SVM classifier to recognize static sign using kinect. Various researchers are working on Arabic sign language recognition for isolated word recognition using various methods such as, pulse coupled neural network (PCNN) [16] , HMM [17], simple KNN [8]. Most of these techniques required large number of training samples and are mostly dependent on the signer. In reality, signer independent method is more practical and desirable. Most of the work is on isolated word recognition where, continuous word and sentence recognition are required for practical use. Complexity of sign language recognition increases from alphabet recognition to word recognition and to continuous sentence recognition. There are many challenges in continuous sign language recognition such as, detecting and modeling extra movements of end of a sign and start of a new one. Until now, researchers are treating this problem as speech recognition problem using HMM classifier [10] and graph matching [11] methods for continuous sign language recognition. Size of vocabulary is a major issue in any SL interpretation for implementation purpose. So, researchers are focusing on few vocabulary signs for testing the result. Continuous sign language recognition is more practical as well as challenging at present.

3. Grammar Model for ISL Dataset

ISL consists of manual as well as non-manual gestures where, manual gesture considers hand parameter such as, hand shape, location, orientation and movement whereas, non-manual gesture consists of facial expression, lip movement and movement of other body part [3]. The scope of this paper is limited to manual gestures as hand plays a very important role in any sign language as well as any HCI applications. Few vocabulary sets of ISL [18]were considered for testing the result. A grammar model was designed using skeleton information obtained from Microsoft Kinect sensor. Table 1 shows the description of ISL word based on skeleton joint information and quadrant information. Motion pattern was created based on this information. Quadrant information is a signer space divided into four parts with reference to X and Y axis.

Here, ISL data set was modeled for implementation purpose using syntactic approach with the help of production rule. All non-terminal symbols are denoted in capital letter. Production rules for ISL dataset are given as follows. Figure 1 shows parse tree generated with given grammar for sign FLY and NAME for illustration.

3.1 Grammar Model for ISL

ISL$\rightarrow$ MG | NMG | MG NMG ▷ MG : Manual gesture, NMG : Non-manual gesture

MG$\rightarrow$ O | T ▷ O: One handed, T: Two handed

Table 1: ISL Word Signs Description for sample cases

Sign	one or two hand	Quadrant location (Initial position, Final position)	Joints involved
Fly	Two Hand	RH 1 RH 4 LH 2 LH 3	Hand, Wrist, Elbow, Shoulder
Name	One Hand	RH X axis	Hand, Wrist, Elbow, Shoulder
Friend	Two Hand	LH 4 LY axis RH 3 RY axis	Hand, Wrist, Elbow, Shoulder
Circle	One Hand	RH 1 RH 1	Hand, Wrist, Elbow, Shoulder
Bye	One Hand	RH 2 RH 2	Hand, Wrist, Elbow
Correct	One Hand	RH 2 RH 1	Hand, Wrist, Elbow, Shoulder
Body	Two Hand	LH 1 LH 4 RH 2 RH 3	Hand, Wrist, Elbow, Shoulder
Sleep	Two Hand	LH 1 LH 1 RH 2 RH 1	Hand, Wrist, Elbow, Shoulder
Sit	Two Hand	LH 1 LH 4 RH 2 RH 3	Hand, Wrist, Elbow, Shoulder
Square	Two Hand	LH 1 LH X axis RH 2 RH X axis	Hand, Wrist, Elbow, Shoulder

O $\rightarrow$ S | D $\triangleright$ S: Static gesture, D : Dynamic gesture
T $\rightarrow$ S | T' D $\triangleright$ T': type 0 and type 1
T'$\rightarrow$ t0 | t1 $\triangleright$ t0 : both hands active, t1 : One hand dominant over other
S $\rightarrow$ SHAPE'
SHAPE' $\rightarrow$ SHAPE SHAPE' | ϵ
SHAPE $\rightarrow$ f1 | f2 | f3| ... |fn $\triangleright$ f1 to fn are feature descriptors
D $\rightarrow$ HJ'D'| ϵ $\triangleright$ HJ: hand skeleton joints,
HJ'$\rightarrow$ HJ HJ'| ϵ
HJ $\rightarrow$ h |w | e | s $\triangleright$ s: shoulder joint, e: elbow joint, w: wrist joint, h: hand joint
D'$\rightarrow$ LQM'| RQM' | LQM'RQM' $\triangleright$ LQM': Quadrant motion w.r.t left, RQM': quadrant motion w.r.t right hand
LQM'$\rightarrow$ LQM motion LQM' | ϵ $\triangleright$ motion: movement of gesture from one location to other
RQM'$\rightarrow$ RQM motion RQM' | ϵ
LQM $\rightarrow$ lh1 | lh2 | lh3 | lh4 | lhx | lhy $\triangleright$ lh: left hand, x: X- axis, y: Y-axis, 1,2,3,4 are quadrant number
RQM$\rightarrow$ rh1 | rh2 | rh3 | rh4| rhx | rhy $\triangleright$ rh: right hand, x: X- axis, y: Y-axis, 1,2,3,4 are quadrant number
NMG' $\rightarrow$ NMG NMG' | ϵ
NMG$\rightarrow$ fe | ez | bp | lm $\triangleright$ fe: facial expression, ez: eye gaze, bp: body posture, lm: lip movements

4. Dynamic Hand Gesture Recognition and Sentence Interpretation

Here, focus has been given to ISL word recognition and sentence interpretation which are the next steps after alphabets and numbers recognition. The methodology consists in dividing

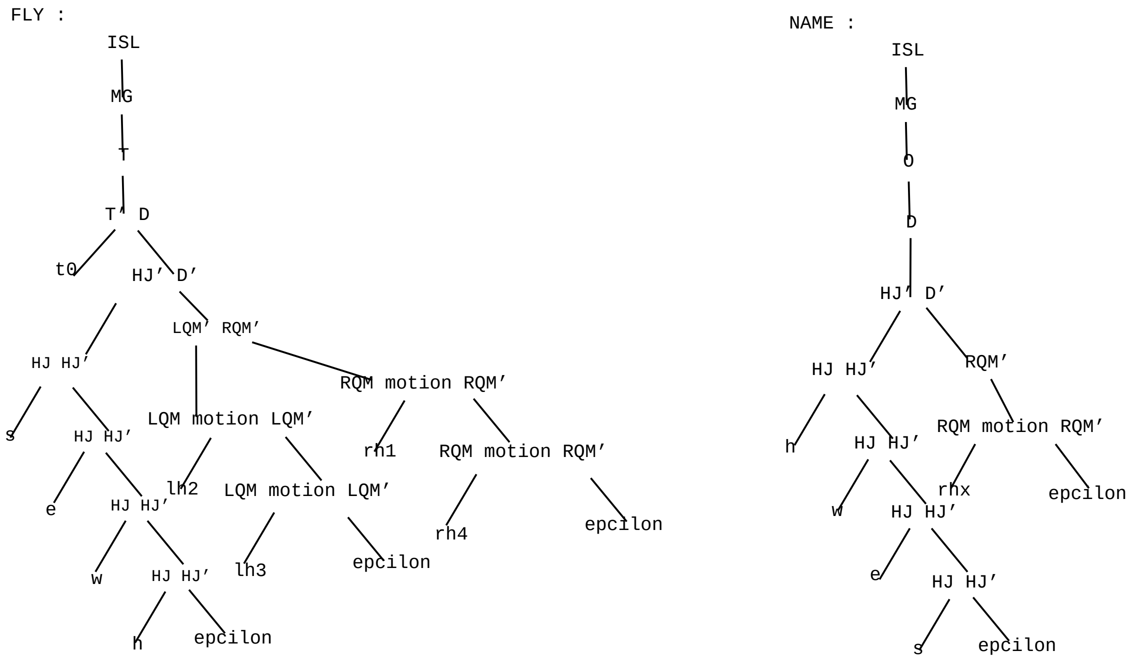

Fig. 1: Parse trees for sign NAME and FLY using Grammar Model

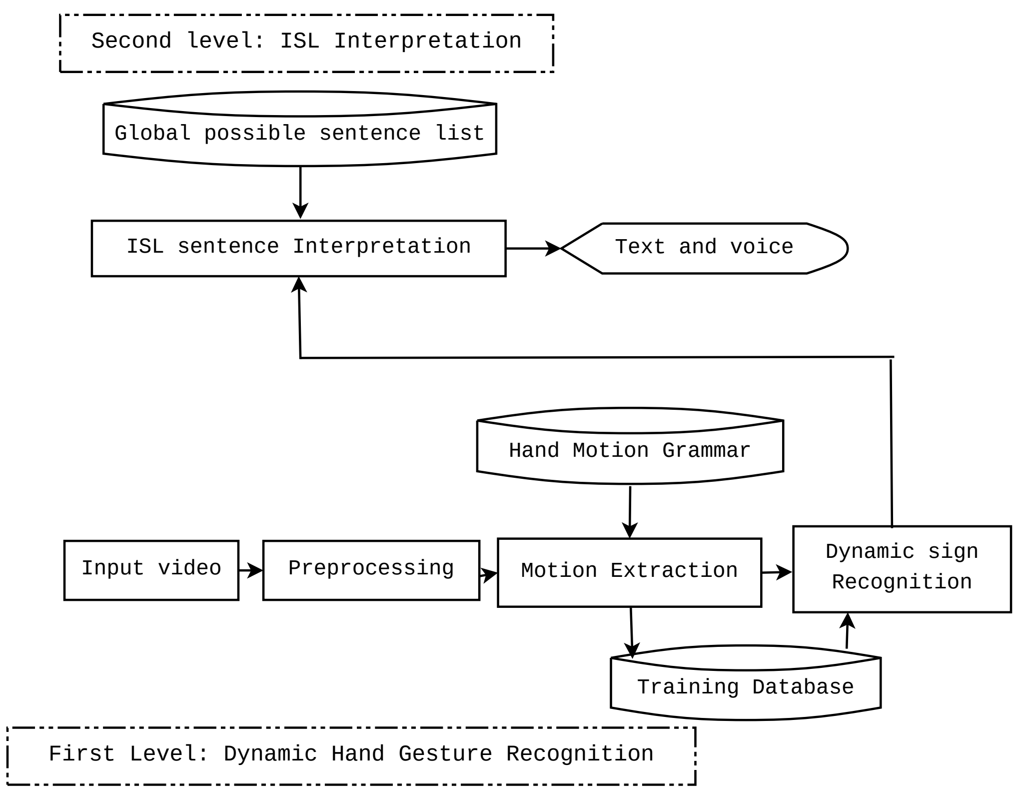

Fig. 2: Two tier Architecture for Dynamic Hand Gesture Recognition

Table 2: Terminology used for gesture description in decision rule based algorithm

Left hand	Right hand	X	Y	Z	Hand joint	Wrist joint	Elbow joint	Shoulder joint
L	R	x	y	z	h	w	e	s

Table 3: Gesture description with skeleton joints for decision rule based approach

(Hand, hand joint, axis)	Description
(Rhx, Rhy, Rhz)	Right hand with hand skeleton joint with reference to X, Y and Z axis
(Rwx, Rwy, Rwz)	Right hand with wrist skeleton joint with reference to X, Y and Z axis
(Rex, Rey, Rez)	Right hand with elbow skeleton joint with reference to X, Y and Z axis
(Rsx, Rsy, Rsz)	Right hand with shoulder skeleton joint with reference to X, Y and Z axis
(Lhx, Lhy, Lhz)	Left hand with hand skeleton joint with reference to X, Y and Z axis
(Lwx, Lwy, Lwz)	Left hand with wrist skeleton joint with reference to X, Y and Z axis
(Lex, Ley, Lez)	Left hand with elbow skeleton joint with reference to X, Y and Z axis
(Lsx, Lsy, Lsz)	Left hand with shoulder skeleton joint with reference to X, Y and Z axis

the task into two levels as described in Figure 2. The first level, involves dynamic hand gesture recognition for ISL words and second one, ISL sentence interpretation.

4.1 Dynamic Hand Gesture Recognition Using Decision Rule Based Approach

ISL dynamic manual sign consists of sequence of postures connected with motion over a time period. So, motion of the hand was tracked and motion descriptor was created with the help of skeleton viewer. Initial position of the signer hand was first set for translation invariance at the time of recognition. Training was not required for this algorithm. Decision rules needed to be defined based on the skeleton joint information. Motion descriptor in the form of (x, y, z (depth)) values for required skeleton joint information, was taken for each hand. Considering initial and final position, decision rules were created and checked at run time for recognition purpose with defined threshold value(T). Table 2 describes the terminology used for gesture description. Table 3 explains the gesture description with its meaning. In table 4 decision rules are given for few signs. While forming the rules C is used for current position, M is used for measured position and T is used as threshold value for motion parameter.

4.2 Dynamic Hand Gesture Recognition Using DTW

In any SL interpretation, major challenge is the speed of signer which, may vary in time domain for same sign from person to person or for the same person. To overcome this challenge, DTW [12] distance measure was used to find similarity between two motion sequences. It calculates an optimal match between two given sequences of feature vectors

Table 4: Decision rules for few signs

Sign	Decision rules
Bank	$(CRhz > (MRhz - T))\&\&(CRhx < (MRhx + T))$
Await	$(CRhz > (MRhz - T))\&\&(CRhz < (MRhz + T))$
Boss	$(CRhy < (MRhy + T))\&\&(CRhy > (MRhy - T))$
	$\&\&(CRhx < (MRhx + T))\&\&(CRhx > (MRhx - T))$
Open	$(CRhx < (MRhx + T))\&\&(CLhx > (MLhx - T))$
Driver	$(CRwy < (MRwy + T))\&\&(CLwy > (MLwy - T))$
	$\&\&(CRwz < (MRwz + T))\&\&(CLwz > (MLwz - T))$

which allows for stretched and compressed section of the sequence. It is a well-known method which has been used in various applications such as, signature analysis and speech recognition. Here, each sign was studied with the parameters such as, skeleton joints involving one/two handed signs. Accordingly, pattern was created in the form of temporal sequence $(x_1, y_1, d_1)....(x_n, y_n, d_n)$ for n number of video frames. Frame rate was considered as 30f/s. Here, each (x_i, y_i, d_i) represents the pixel coordinate and depth information of joint in each frame. Single sample training was required for each sign and was stored in database. The algorithm 1 and 2 are presented for ISL word recognition using DTW. Algorithm 1 is used to store single pattern per sign into database mentioned as Plist. Algorithm 2 is used to test unknown sign using DTW distance measure with stored pattern.

Algorithm 1 Dynamic Hand Gesture Recognition Training algorithm using DTW

▷ Training is for 1 to N sign gestures and stored into PatternList Plist
▷ PatternList Plist= $\sum_{i=0}^{N} Plentry_i$; N number of pattern stored into list
procedure TrainGesture
  PatternList Plist; ▷ Plist is a list of trained patterns.
  for i=1 to N **do**
    j=$GetSkeletonJoints()$;
    $\forall frames$ ▷ For all video frames process following steps
    for k=1 to j **do**
      $FeatureVector_k = ExtractFeature(k)$ ▷ Add a skeleton time sequence with a label to pattern list
    end for
AddInPatternlist(Plist,FeatureVector,label)
  end for
  return Plist;
end procedure

Algorithm 2 Dynamic Hand Gesture Recognition using DTW

▷ $\theta 1$ is the Maximum distance between the last observations of each train sequence with current test sequence
▷ $\theta 2$ is the Maximum DTW distance between a test example and a train sequence being classified
procedure RecognizeGesture(PatternList Plist, InputTestSequence)
  minDist = 9999;
  classification = "UNKNOWN";
  for i=1 to N **do** ▷ Number of trained gesture pattern in Plist
    for j=1 to M **do** ▷ Number of features of each pattern
      if (CalculatedTestPositionDistance(InputTestSequence,Plist)$< \theta 1$) **then**
        $d = DTW(InputTestSequence, Plist_i)$
      end if
    end for ▷ calculate distance using DTW
    if $d < minDist$ **then**
      $minDist = d$;
      $classification = label$;
    end if
  end for
  return ($minDist < \theta 2$? classification : "UNKNOWN")
end procedure

4.3 ISL Sentence Interpretation

Linguistic study, psychology and culture play a very important role in the formation of sentence. Considering all these things, SL interpretation using vision and pattern recognition system is a great challenge for researchers. Apart from that, there is a need to distinguish a static sign from a dynamic sign or an accidental hand posture [11].

4.3.1 Challenges In Sign Language For Word And Sentence Recognition

Sign language function and spoken language functions are totally different. SL is fundamentally based on spatial properties. Hand parameters such as shape, movement, orientation and location as well as facial expression, lip movements are considered to interpret the sign. These parameters occur simultaneously and are articulated in space. Building of syntax and semantics based rule system is required because one sentence in a spoken language can be represented by a single sign in SL. Britta and Karl-Friedrich [21] reported some difficulties in respect of sign language in general: i) occlusion problem while performing sign ii) signer position may vary in front of camera while performing a sign iii) working with 2D camera led to loss of depth information iv) as each sign varies in time and space, so there may be a change in position and speed with the same person or from person to person v) co-articulation problem (link between preceding and subsequent sign). Vishwanathan and Idicula [22] discussed that ISL differs in the phonology, morphology, grammar and in syntax from other country languages. ISL signs are more complex compared to other SL signs. It consists of both handed complex and ambiguous signs. For some signs, hand contact to other body part is required. Local and global movements are required for some signs. For construction of the sentence spoken language and sign languages are totally different in their linguistic structure. In ISL, proper noun is not pronounced as in spoken language. It is used in reference to his/her gender (male/female) followed by physical description of the person referred to. The structure of sentence in English language is SUBJECT, VERB, OBJECT. But in ISL, this structure is not always used. In some cases, it is SUBJECT, OBJECT, VERB [19]. These are some of the important linguistic properties which make ISL interpretation a challenging task. Sign language is as complex as any spoken language. It is having thousands of signs, each differing from next by minor changes in hand shape, motion, position, non-manual features or context. Signs can be described at the sub-unit level using phonemes. These encode different elements of signs. Unlike speech they do not have to occur sequentially, but can be combined in parallel to describe a sign [22] . There are some significant challenges in the field of SLR such as [20]

- Adverb modifying verb: signer would not use two signs for run slowly/quickly. They would modify run sign with speed.
- None manual features: facial expression and body posture are key in determining the meaning of sentences, e.g. eyebrow position, lip movement can determine the meaning with manual sign.
- Pronoun: such as he, she, or it do not have their own sign, instead, the referent is described and allocated a position in the signing space.
- Directional verbs: the direction of the motion describes the direction of the verb e.g. give and pen
- Positional signs: describes body part where the sign acts.

Table 5: Input keywords and possible sentence for ISL sentence algorithm

Sentence No.	Keywords (ISL signs)	Possible ISL (sentence list)
S1	Correct	Correct! or This is correct.
S2	Wrong	Wrong! or This is wrong.
S3	you, correct	You are correct.
S4	how,you	How are you ?
S5	how,old,you	How old are you ?
S6	Bird, fly	Bird is flying.
S7	You, fly	when are you flying?

- Finger spelling: sign is not known to signer and recipient and spelled by local spoken word by finger spelling, example, can be name of the person, city. SLR problem becomes more challenging with signing style e.g. every signer has his own style, as also with handedness of the signer who may be left handed or right handed.

4.3.2 ISL Sentence Interpretation Using Inverted Indexing

As discussed difficulties for SL interpretation,to recognize each and every word with grammar and convert it into sentence is really a challenging task. So, intelligent solution with inverted indexing concept is presented, which is divided into two parts i) creation of index table and ii) searching of possible sentence based on input keyword using index table. Continuous sentence recognition for any SL has major challenges for vision based system which is discussed earlier. The algorithm considered only important recognized keywords of a sentence and interpreted meaningful sentence using inverted indexing concept. The inverted indexing concept is mostly used in the field of information retrieval such as, for search engine indexing algorithm as well as in bio-informatics and in general, document search algorithm [23]. The proposed work is identified this method for use of sentence interpretation in SL. Recognized signs as keywords and possible sentence which are described in table 5 were given as an input to algorithm 3. Each keyword was stored along with sentence number list in index table. For each insertion of keyword,index table was scanned and corresponding entry was updated. The detailed steps are given in algorithm 3 which creates index table 6. Once the index table is created, next step is finding the possible sentence based on input keywords. To illustrate the proposed algorithm, let us take one example. Consider K is the list of recognized input sign words given below. After finding sentence no. against each input keyword ,intersection operation was applied on each sentence no. list and final sentence no. was retrieved and corresponding sentence was displayed. Algorithm 4 gives output FS_n as a sentence number.

Table 6: Index table generated for Table 5 inputs

Keyword ISL word	Sentence No.
correct	S1,S3
Wrong	S2
You	S3,S4,S5,S7
How	S4,S5
Old	S5
Bird	S6
Fly	S6,S7

Algorithm 3 Index Table Construction for ISL Sentence Interpretation Algorithm

▷ Input: Input data set
▷ $E = \{Sn, Kl, S\}$ Each entry in dataset D consists of sentence no, keyword list and sentence
▷ $D := \sum_{i=0}^{N} E_i$; input dataset for N sentences
▷ Output: Index Table
▷ $IndexTableEntry = K, Sl$; Each index table entry consists of keyword and sentence list
▷ $IT = \sum_{j=0}^{M} I_{TEi}$ ▷ M entries in index table
▷ SKS is stored keyword list into index table

procedure CONSTRUCTINDEX(Dataset D)
  SKS=ϕ; index,i,j=0; ▷ Initially stored keyword list is empty
  for i=0 to N **do** ▷ N no. of entries in D
    for each K in keyword list **do** ▷ Construct index table for each keyword
      if $K \epsilon SKS$ **then**
        Index= GetIndex (K); ▷ If keyword is already present get index of it.
        $IT_{Index}.Sl = \cup D_i.Sn$; ▷ Append new sentence no. with existing sentence no. list.
        else
        $IT_j.K = K$; ▷ If new keyword then insert keyword and sentence no. with new entry in index table.
        $IT_j.Sl = D_i.Sn$;
        $SKS_j = K$;
        j++;
      end if
    end for
  end for
  return IT;
end procedure

Algorithm 4 Sentence Interpretation Algorithm using Index Table

▷ Input : Index Table IT, Input keywords;
▷ Output: Output set O ; set of sentence no. list and final statement no. F_{Sn}
procedure SentenceInterpretation (IT,InputKewords)
  Index,i=0;
  for each K in InputKeywords **do**
    $Index = GetIndex(K)$; ▷ Get index for input keyword
    $O_i = IT_{Index}.Sl$; ▷ Get sentence no. list for each keyword
    i++;
  end for
  $F_{Sn} = (O_0 \cap O_1..\cap O_n)$; ▷ By applying intersection operation find correct possible sentence no.
  Display (F_{Sn});
end procedure

suppose K= how,old,you ▷ Input recognized signs as a keywords
IT={S4∪ S5},{S5},{S3∪ S4∪ S5 ∪ S7} ▷ Output of Algorithm 3
FS_n={S4,S5}∩ {S5}∩ {S3,S4,S5,S7} ▷ Output of Algorithm 4
FS_n= {S5, How old are you?} ▷ Possible sentence is displayed

5. Experiment Results

This section describes the result of each proposed algorithm.

5.1 Result of Dynamic Hand Gesture Recognition Using Rule Based Algorithm

Decision rule based approach was working well for few subset of ISL vocabulary, however, for the two closest sign it gave misclassification. Figure 3 shows result for the recognized signs "BOSS" and "CAR" as a sample case. This was a user independent algorithm and working well for up to 7 vocabulary signs. Here, misclassification rate increases with increase in vocabulary. This is a very good algorithm for any HCI application where, small set of signs is required such as, multimedia control and desktop control applications.

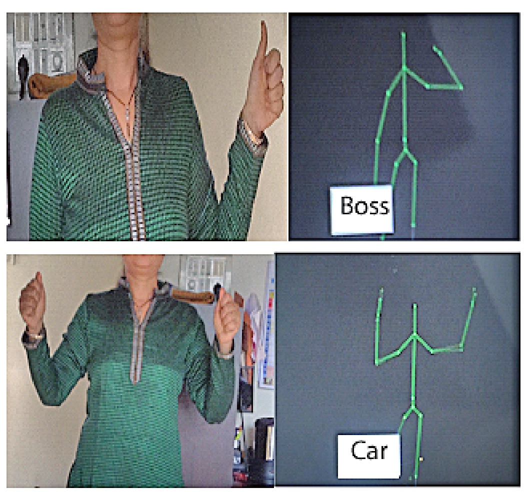

Fig. 3: Experimental sample sign words result using rule based approach for "boss" and "car"

Table 7: Confusion matrix for ISL word recognition using DTW for few samples

Recognition accuracy(%)	name	circle	bye	body	happy	lazy	sign-language	you	me	bank	unknown
name	84										16
circle		90									10
bye			98								2
body				84	16						
happy				10	90						
lazy						80					20
sign-language	10						90				
you								90	10		
me								10	90		
bank										98	2

5.2 Result of Dynamic Hand Gesture Recognition Using DTW

In this work, initially few ISL signs were trained and tested using DTW distance measure. It was user-independent algorithm where, only one sample was required to train per sign. Testing was done on five different persons on 10 attempts (N) for each sign. Accurate result was achieved for expert signer. For non-expert signers, in first two attempts signs could not be correctly recognized due to the vast variation in speed. For any SL, identification of size of vocabulary is a challenge for practical implementation. Work is in progress for adding more vocabulary. Recognition rate was found to be high for expert signer and average recognition rate of 96.25% was achieved for 40 signs. Algorithm was working for one as well as two handed signs. Table 7 shows confusion matrix for few important sample cases. The signs that gave 100% accuracy are not mentioned to reduce size of the table. In most of the sign, if test sign not matched with stored sign then it was detected as a unknown sign. This was helped us to reduce false positive rate. Each sign for each person tested for 10 times and average accuracy was calculated using equation 1 for few ISL words.

$$Average\ recognition\ accuracy(\%) = \sum_{i=1}^{n}\sum_{i=1}^{m}\frac{Correct\ recognition}{nm} \times 100 \tag{1}$$

here, n is number of different signers and m is number of signs

The proposed work is compared with related ISL work. There is a lot of research being done on various foreign sign languages. Comparing the proposed work with other SL is not feasible and appropriate due to difference in morphology, phonology and grammar of each SL. However, reported work on other SL is included in literature survey. In India, on standard English ISL still, very less work has been done for dynamic signs. Nandy et al. [24] used ISL for human robot interaction. They have used corpus of 9 words and achieved 90% accuracy using Euclidean distance. Kishore and Kumar [13] worked on video based ISL isolated sign recognition using Fuzzy logic and achieved 96% accuracy. Proposed approach used DTW for continuous sign recognition method with skeleton joint depth features and achieved 96.25% accuracy.however, results may vary with different vocabulary and size of vocabulary. In reported work, authors are worked on isolated sign recognition where proposed work presented for continuous dynamic sign recognition.

5.3 Result of ISL Sentence Interpretation Using Inverted Indexing

Figure 4 shows the result generated by the ISL sentence algorithm. With few recognized words possible sentence can be interpreted which overcome the drawback of traditional

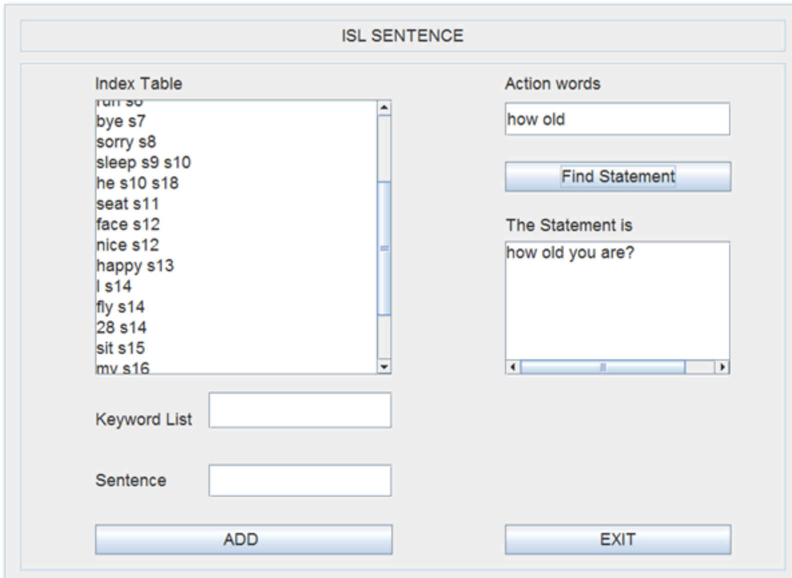

Fig. 4: Experimental results for ISL sentence creation using inverted indexing concept

Table 8: ISL Sentence Interpretation Results

Recognized sign words	Possible ISL sentence given by Algorithm 3 and 4
(sit)	Sit down!
(Face, Nice)	You are looking nice
(I ,fly,28)	I am flying on 28th
(you, happy)	You are looking happy
(Where, you, born)	Where were you born?
(People, house, live, how)	How many people live in house
(You, correct)	You are correct
(I, happy)	I am happy
(You, lazy, ?)	Are you lazy?
(You, lazy)	You are lazy.
(Wonderful)	This is wonderful.
(How, old, you)	How old you are?
(Push, door)	Push the door
(Pull, door)	Pull the door
(My, fight, you)	My fight is with you
(I, Small, house)	This is a circle
(I, correct)	I am correct
(He, waiting, afternoon)	He was waiting for you in the afternoon
(India, more, deaf and dumb, world, country)	There are more deaf and dumb people in India as compared to other countries in the world
(World, sign language, different country)	In the world, each country is having different sign language

sentence interpretation method where each sign with required grammar need to be recognized.Table 8 describes few sample cases out of 100 test samples of ISL interpretation result.

6. Conclusion and Future Work

Comparison analysis of proposed two algorithms for dynamic hand gesture recognition revealed that, with any approach, increased vocabulary recognition rate decreases. However, DTW based approach gave better recognition accuracy with more vocabulary than rule based approach. The proposed ISL sentence creation algorithm using inverted indexing concept showed simplified solution for sentence interpretation problem and it overcame few of the challenges of continuous sentence recognition such as neglecting grammar symbol and sentence structure part. The major advantage of this approach was ISL interpretation system which could interpret meaningful sentence with few input recognized words. However, what was needed was that a sentence was interpreted according to the possible sentence list and keyword that were stored. Sometimes exact sentence might not have been interpreted but thoughts having same meaning were conveyed. Currently 100 possible sentences are tested. Accuracy of ISL sentence interpretation was 100% if input keyword and sentences were properly trained. Using inverted indexing concept for SL interpretation is a novel approach yet to be explored. This simple and optimized solution can be used in any SL interpretation to overcome the challenges of continuous sentence recognition. Further work will be carried out to increase vocabulary and sentences. Dynamic hand gesture recognition algorithm will be further modified using more features to make a system robust and accurate with increased vocabulary. Modified DTW algorithm will be developed where, training dataset will be separated on the basis of single and two handed sign to increase accuracy and decrease misclassification rate.

References

[1] ISLRTC *ISLRTC-Quarterly newsletter, Indian Sign Language Research and Training Centre.* IGNOU, 2012.

[2] U. Zeshan, M. Vasishta, M. Sethna, *Implementation of Indian sign language in educational setting.* Asia Pacific Disability Rehabilitation Journal, 16(1), pp.16-39, 2005.

[3] T. Dasgupta, S. Shulka, S.Kumar, S. Diwakar, A. Basu, *A Multilingual Multimedia Indian Sign Language Dictionary Tool.* Proceedings 6th Workshop on Asian Language Resources,pp.57-64, 2008.

[4] J. Han,L. Shao,D. Xu,and J.Shotton *Enhanced Computer Vision with Microsoft Kinect Sensor : A Review.* IEEE transaction on Cybernetics,43(5), pp.1318-1334, 2013.

[5] C. Vogler,D. Metaxas, *A Framework for Recognizing the Simultaneous Aspects of American Sign Language.* Computer Vision and Image Understanding,pp.358-384, 2001.

[6] Y. Sun,N. Kuwahara, and K. Morimoto, *Analysis of recognition system of Japanese sign language using 3D image sensor.* IASDR,pp.1-7, 2013.

[7] K. Stefanov, and J.Beskow, *A kinect corpus of Swedish sign language Signs.* Proceedings Workshop on Multimodal Corporation,pp.1-5, 2013.

[8] T. Shanableh,and K. Assaleh, *Arabic sign language recognition in user-independent mode.* In Proc. int. Conf. Intell. Adv. Syst,pp.597-600,2007.

[9] A. Kuznetsova,L. Leal-Taixe, and B. Rosenhahn, *Real-time sign language recognition using consumer depth camera.* ICCV , IEEE Xplore,pp.83-90,2013.

[10] K. Assleh, T.Shanableh,M.Fanaswala, F. Amin, and H. Bajaj, *Continuous Arabic sign language recognition in user dependent mode.* J. Intell. Learn. Syst. Appl., 2(1), pp. 19-27, 2010.

[11] M. F.Tolba, A. Samir, and M. Aboul-Ela, *Arabic sign language continuous sentences recognition using PCNN and graph matching.* Neural Computing Application, 23(3), pp.999-1010, 2013.

[12] J. Lichenauer, E. Hendriks, and M. Reinders, *Sign Language Recognition by Combining Statistical DTW and Independent Classification* IEEE Transaction on Pattern Analysis and Machine Intelligence, 30(11), pp. 2040-2046, 2008.

[13] P. Kishore,P. Rajesh Kumar,E. Kiran Kumar, and S.Kishore, *Video Audio Interface for Recognizing Gestures of Indian Sign Language.* International Journal of Image Processing (IJIP),5(4), pp. 479-503, 2011.

[14] F. Huang, and S. Huang, *Interpreting American Sign Language with Kinect* pp.1-5, 2011.

[15] Z.Hu,L.Yang,L. Luo,Y. Zhang, and X.Zhou, *The Research and Application of SURF Algorithm Based on Feature Point Selection Algorithm.* Sensor and Transducers, IFSA publishing, pp.67-72, 2014.

[16] A. S.Elons, M. Abull-Ela, and M.F.Tolba, *A Proposed PCNN Features Quality Optimization Technique for Pose-invariant 3D Arabic Sign Language Recognition.* International Journal of Knowledge-Based Intelligent Engineering System, vol.14, No.3, pp.139-152, 2010.

[17] M. Mohandes,and M. Deriche, *Image based Arabic Sign Language Recognition.* In Proceeding 8th Int. Symp. Signal, vol.1, pp.86-89, 2005.

[18] *Instructional Indian sign language video: A project of International human resource development centre (IHRDC) for the disabled.* Ramkrishna mission vidyalaya, Coimbatore. http://indiansignlanguage.org.

[19] A. Ghotkar and G.Kharate, *Study of Hand Gesture Recognition for Indian Sign Language.* International Journal of Smart Sensing and Intelligent Systems, Vol. 7, No.1, pp. 96-115, March 2014.

[20] H. Cooper, B. Holt, R. Bowden, *Sign language recognition.* eprint13.blaclnight.ie/531441/1/SLR.LAP.pdf, pp.1-20, 2013.

[21] B.Bauer , Karl-Friedrich *Towards an Automatic Sign Language Recognition System Using Subunits.* LNAI 2298, GW-2001,Springer pp.34-47,2001.

[22] D.Vishwanathan and S. Idicula *Recents Developments in Indian Sign Language Recogniton: An Anaylysis.* International Journal of computer Science and Technologies Vol.6,No.1, pp.289-293, 2015.

[23] J. Zobel, and A. Moffat, *Inverted Files for Text Search Engines.* ACM Computing Surveys, Vol.38,No.2, pp.1-56, 2006.

[24] A. Nandy, S. Mondal, J. Shankar, P. Chakraborty and G. Nandy, *Recognizing and Interpreting Indian Sign Language Gesture for Human Robot Interaction.* International Conference on Computer and Communication Technology, pp. 712-717,2010.