

Speech to Gesture Interpreter

S. S. Ashwini¹, Niharika², A. Vyshnavi³

¹Assistant Professor, Department of Information Science Engineering, Nagarjuna College of Engineering and Technology, Bangalore, India

^{2,3}Student, Department of Information Science Engineering, Nagarjuna College of Engineering and Technology, Bangalore, India

Abstract: Sign language is a natural way of communication for who are with hearing and speaking disabilities. There have been various techniques developed to recognize and convert this sign language to the understandable text or vice-versa. The projects that have been so far developed are either speech to text converter or text to sign converter. Our project aims at creating a translation system that converts the input speech to the human understandable sign language.

Keywords: Speech to Gesture.

1. Introduction

The data available is mostly in the form of written or spoken language which are dependent on different lifestyles and preferences of people. So, the general language and the sign language used for communication are different. Previous work on sign language translation has mostly defined one-to-one relations between unit expressions in spoken and sign languages, along with simple translation rules. However, they suffer from the problem of not really reflecting the differences in the information space that the two types of language have. It is thus not surprising that an expression in spoken language becomes either unnecessarily verbose or severely restricted in its counterpart in sign language.

2. Related work

A. Sign language vs. Natural language

Sign language is not just the medium for substituting other natural languages, but is also uniquely specialized for the representation of thoughts and expressing them in understandable form. Sign languages use human hands and their fingers for various movements and actions, and use the human body, forming a vision-movement system, which can be contrasted from the audio-vocal system of spoken languages. The basic units of information for sign language are hand-shape, hand-position, hand-motion and hand-direction, as opposed to phonemes in spoken languages [1]. There are several sign languages available such as American, British, German, French, Italian, and Turkish Sign Language. American Sign Language (ASL) is well-known and the best studied sign language in the world. The grammar of ASL has been applied to other sign languages especially as in British Sign Language

(BSL). The first real study of sign languages is achieved in 1960s. Dr. William C. Stokoe published the monograph Sign Language Structure [2] in 1960.

3. Existing system and its drawbacks

A. Existing System

Although sign language is used across the world to bridge the gap of communication for hearing or speech impaired which depend mostly on sign language for day to day communication, there are not efficient models that convert text to Indian sign language. There is a lack of proper and effective audio visual support for oral communication. While significant progress has already been made in computer recognition of sign languages of other countries but a very limited work has been done in ISL computerization.

Work done so far in this field has been much more focused on American sign language (ASL) or British sign language, but for Indian sign language, systems that have been developed are very few. The underlying architecture for most of the systems are based on:

1. *Direct translation:* Words from source language are directly transformed into target language words. Output may not be a desired one.
2. *Statistical Machine Translation:* It requires large parallel corpus, which is not readily available in case of sign language.
3. *Transfer based architecture:* Grammar rules are being applied so as to define a proper translation from one language system to another.

Another is "Sign to Text Converter", the main objective is to translate sign language to text. The framework provides a helping-hand for speech-impaired to communicate with the rest of the world using sign language. This leads to the elimination of the middle person who generally acts as a medium of translation. This would contain a user-friendly environment for the user by providing text output for a sign gesture input [1].

Disadvantages of Existing System:

- The system is not reliable as it does not count the special symbols such as bullets, dashes etc.
- Text may change as pronunciation changes.
- Homonyms are difficult to convert to text.

B. Problem Statement

Sign language is a language that uses manual communication methods such as facial expressions, hand gestures and bodily movements to convey information. This project makes use of videos for specific words combined to translate the text language into sign language.

4. Proposed system and its advantages

Few works have been done to generate a system that is based on the above concepts listed in the existing approaches section and cater to Indian sign language. Thus we propose to develop one for Indian sign language based on transfer based translation.

The success of this translation system will depend on the conversion of English text to Indian sign language bearing its lexical and syntactic knowledge. As mentioned in existing system the disadvantages can be overcome by developing a system which is "Audio to Sign Language Converter". This system proposes a solution to this problem by providing an audio to sign language translation. Speech processing which includes Speech Recognition is the study of recognizing the words being spoken, regardless of who the speaker is.

This project is based on converting the audio signals received to text using speech to text API (python modules or Google API) and then using the semantics of Natural Language Processing to breakdown the text into smaller understandable pieces which requires Machine Learning as a part. Data sets of predefined sign language are used as the input so that the software can use artificial Intelligence to display the converted audio into the sign language.

Advantages of proposed system:

- Smart algorithms are being used to increase the performance and for accuracy.
- Can be used for any type of special symbols.
- No need of any translator medium to translate.

5. Interlingua based translation

In this system, a language independent semantic structure is produced by doing only semantic analysis on the original input text. This independent structure is called Interlingua. The target language is then generated from this Interlingua. It is thus an alternative of both direct and transfer based translation.

A. Facts about sign language

Sign languages is not just 'natural language represented by sign' or not just hand representation of the words as it is but rather it is the representation of meaning. There are various facts associated with sign language, a natural language, which most of us are unaware [2]. Some of them are listed below:

- NOT the same all over the world
- NOT just gestures and pantomime but do have their own grammar.
- Dictionary is smaller compared to other languages.

- Finger-spelling for unknown words.
- Adjectives are placed after the noun for most of the sign language.
- Never use suffixes
- Always sign in present tense
- Do not use articles.
- Do not use I but uses me.
- Have no gerunds
- Use of eyebrows and non-manual expressions.
- NOT been invented by hearing people.

6. Semantic representation

The approach to English to Sign Language translation is based upon use of Discourse Representation Theory (DRT) (Kamp Reyle 93) for the intermediate representation of meaning. A DRS (Discourse Representation Structure) is a two-part construction involving a list of variables denoting the nominal discourse referents and conditions (a collection of propositions which capture the semantics of the discourse). DRT was chosen as the underlying theory because it decomposes linguistic phenomena into atomic meaning components (propositions with arguments), and hence allows isolation of tense/ aspect and anaphoric connections that are realized in different sign language grammatical constructs or modalities [3].

DRSs, described in (Kamp Reyle 93), are modified to achieve a more sign language oriented representation that subsequently supports an easier mapping into a sign language grammar. In (Kamp Reyle 93) only event propositions are labeled for use as arguments with temporal predicates. This has been extended by introducing labels for each semantic predicate. As in Verbmobil the labeling of all semantic entities allows a flat representation of the hierarchical structure of arguments and operator embeddings (VIT representation) (Dorna Emele 96; Dorna 00). The label arguments referring to other elements in the flat list are useful to handle verb modifiers, negation and adverbials (e.g.: [attr1: big(X), attr2: very(attr1)]), which reflect the multiple modalities in sign languages: facial expressions convey intensity, head nod (parallel to signing) negation. This form of representation also has the advantage, as Dorna Emele claim, that additional constraints which are important for generation in the target language, e.g. topic/focus in sign languages, may be made explicit. Additionally, in contrast to Verbmobil's uniform labeling an ontology for all DRS propositions has been introduced to ease the mapping between the flat semantic structure of the DRS to the input structure of the target language specific HPSG, as required by the generation algorithm in ALE (Carpenter Penn 99). The class definitions, that result from the taxonomy of labels, can help with the search of possible anaphoric referents and temporal relationships.

7. Conclusion

This paper presented an overview on speech to gesture interpreter.

References

- [1] P Blackburn and J. Bos. Representation and Inference for Natural Language. A First Course in Computational Semantics. Volume 2.
- [2] T Veale, A Conway, and B Collins. The challenges of cross-model translation: English-to-sign-language translation in the zardo system. Machine Translation, 1998.
- [3] Goyal, L., & Goyal, V. (n.d.). Automatic Translation of English Text to Indian Sign Language.
- [4] Eom, M. S. 1996. Syntactic analysis of Korean sign languages. Master's thesis, Daegu University, Korea. Davis, H. Silverman, R. 1978. *Hearing and Deafness*. N.Y. Holt Rinehart and Winston.