Native Language Audio To Sign Language Generator

K Chanikya Reddy, A Rakesh, D Anusha, Dr.S.Ramani

Department of Electronics and Communication Engineering,
Sreenidhi Institute of Science and Technology, Hyderabad

Abstract

For those with communication challenges who also have hearing and speaking impairments, sign language comes naturally. The lack of any sign language corpus has prevented the development of many text to sign language conversion systems, despite a variety of tools being available to interpret or recognize sign language and convert it to text. Our project seeks to develop a translation system with a parsing module that converts English sentence structure grammar representation into an Indian sign language grammar representation.

To accomplish this, stop words in the rearranged sentence are removed. Since Indian sign language does not provide word inflections, stemming is used to reduce words to their basic form. The dictionary's word entries with videos for each word are then compared to all the terms in the sentence. If the word is not included in the dictionary, a synonym is used in its stead. The proposed system is novel because it seeks to translate these sentences into Indian sign language grammar in a real-world context, unlike existing systems that just directly translate words into Indian sign language.

Keywords: National language processing, translation, real-world context, voice recognition

Introduction

The research aims to improve the communication ability of those with hearing impairments by translating English text into Indian Sign Language using natural language processing. For informational purposes, sign language makes use of manual communication techniques such facial expressions, hand gestures, and physical movements. With the help of movies for particular words, this project translates text into sign language.

There are no effective models that translate text into Indian sign language, despite the fact that sign language is utilised all over the world to bridge the communication gap for hearing or speech challenged people, who rely primarily on sign language for day-to-day communication. For oral communication, there is a lack of appropriate and efficient audio visual accompaniment. Although there has been great advancement in the computerization of other sign languages, there has been very little done with ISL.

British sign language or American sign language (ASL) have received a lot more attention in this field's work than Indian sign language, where just a small number of systems have been created.

There haven't been many attempts to create an Indian sign language-specific system that is based on the ideas mentioned above and stated in the section on existing techniques. So, we suggest creating one based on transfer-based translation for Indian sign language. The conversion of English text into Indian sign language with understanding of its lexicon and syntax will determine whether this translation method is successful.

Our research seeks to cover the vocabulary of Indian Sign Language, which has about 1500 terms in total. The matching video for each of these words will be gathered. Words that don't fit in this dictionary will be substituted with their synonyms, taking into account word duplication and word parts of speech. If the grammar rules of the two languages disagree, translating one spoken language into another is a difficult undertaking. When the source language is spoken and the target language is sign language, the complexity is greatly increased. The target market for this system is not just hearing-impaired people because it can be very challenging for hearing-impaired people to communicate in public settings like banks, hospitals, train stations, and bus stops because a vocal person may not understand sign language and thus be unable to do so. In order to improve communication, it is therefore directed at everyone who wants to learn this translation.

Literature review

In India, there are 63 million persons, or 6.3% of the population, who have hearing impairments, according to the 2011 census. Among them, 76–89% of the hearing-impaired Indian population is illiterate in spoken, signed, or written languages. The absence of sign language interpreters, the lack of Indian Sign Language tools, or the paucity of Indian Sign Language study are the causes of this low literacy rate. Despite the sophistication of current computers, there has not been much study done on creating MT systems for sign language, especially in India. Many MT methods are employed for other sign languages.

1.System for Direct Translation.

The foundation of this method is word-to-word translation. Sentences' context and meaning are not taken into account. Direct transformation into the target sign language without any changes to the grammar. The original text is not syntactically analysed, and even the conversion order is untouched. Nevertheless, the word order of ISL may not be the same as the given text when translating English text to ISL. To solve this issue, a system that is proficient in both the source and the target languages is required.

2.Translation Based on Transfer

With this system, input is provided as plain text, which is then translated into sign language through syntactic and semantic changes. In this system, the source language is converted into some intermediate text, which is subsequently converted into the target language by using some linguistic rules. It is additionally known as rule-based translation.

3. Translation based on Interlingua

This method uses just semantic analysis on the original input text to create a language independent semantic structure. The name of this autonomous structure is Interlingua. This Interlingua is then used to produce the target language. Hence, it serves as a complement to both direct and transfer-based translation.

Grammar of Indian Sign Language

The manual representation of spoken Hindi or English does not accurately reflect this language's grammar. It has a few distinctive qualities, including:

1. Hand gestures for each hand are used to represent numbers.

2. Male or female prefixes signs of familial connections.

3. All of the WH questions are positioned at the end of interrogative sentences.

4. It also includes a variety of non-verbal gestures, such as eye and head gazing, body postures, mouth motions, and mouth patterns.

5. The symbols for before, then, and after stand in for the past, present, and future tenses, respectively.

Our Approach to the Issue:

By giving hearing-impaired people a video to understand the text's message, the translation module will make it quick and simple for them to comprehend.

Where Our Approach Differs:

Our approach is unique because we transmit information via videos rather than graphics as other solutions do. Moreover, stopwords have been eliminated, which reduces the amount of text processing required.

Hypothesis

A bilingual dictionary is necessary for translation from the source language to the target language. As a result, we'll be making a dictionary that includes both the English term and the Hindu gesture that corresponds to it. The sign language equivalent of the English term could be presented as visuals, videos, or coded text (gloss). All of the methods have advantages and disadvantages of their own, but the video method is the best because it has an advantage over other systems.

Methodologies

Data Gathering

We will download the video clips of each word from http://www.indiansignlanguage.org/. Each video will be manually labelled, and any irrelevant ones will be removed. We would prefer to keep an unfiltered input that includes a variety of words. Programming Design

There are 5 modules in the system:

1. An English parser to analyse texts in English

2. A module for rearrangement of sentences based on ISL grammar rules

3. A stop word eradication tool

4. Stemming, which is used to find a word's roots and substitute synonyms for words that are not in dictionaries.

5. A module for video conversion.

A written English text is provided as the system's input, and after it has been parsed and its grammar represented, a phrase structure is generated. Finally, the text is rearranged to fulfil ISL grammatical requirements since ISL follows Subject-Object-Verb structure coupled with a variety of negative and interrogative sentences, whereas English texts follow Subject-Verb-Object structure. Unwanted words are then eliminated because ISL only uses words that have meanings and does not utilise any linking verbs, articles, or other helpful terms. Whose output is forwarded to the lemmatization module, which breaks down each word into its root form. Synonyms are used in place of words that are not in the dictionary.

<u>**Tools/Libraries Used**</u>

1- Numpy

2- JASigning - Sigml file conversion into animated format

3- movies.py

4- nltk

1. Parsing the English Text Provided

Grammatical structure of both the source language and the target language must be known in order to perform rule-based translation of one language to another. To learn this grammatical structure, parsing is the solution. Three distinct outputs are possible from the Stanford parser: part-of-speech-tagged text, context-free grammar representation of phrase structure, and type dependency representation. When processing an English sentence, the parser employs Penn tree tags.

2. English to ISL conversion grammar rules

As spoken and sign languages each have unique grammar rules. They become considerably more difficult to translate. Hence, according to ISL grammar rules, each of the 20 verb patterns must be used after the occurrence of the associated noun.

3. Removal of Stop Words

Unwanted words are eliminated because ISL focuses on words with specific meanings. These parts of speech include TO, POS (possessive ending), MD (modals), FW (foreign word), CC (coordinating conjunction), some DT (determiners like a, an, and the), JJR, JJS (adjectives, comparative and superlative), NNS, NNPS (nouns plural, proper plural), RP (particles), and SYM (symbols), Interjections and irregular verbs.

4. Synonym substitution and lemmatization

In phrases made in Indian sign language, root words are used. So, we use Porter Stemmer principles to transform them into root form. Also, each word is examined in a bilingual dictionary; if a match cannot be made, a synonym with the same part of speech is assigned.

5.Stage of video conversion

The computer will next look for matches from the dataset that is available for each word once the aforementioned stages have been completed and we have the ISL transformed text. This will be based on the fundamental string matching method between the labels of the movies and the processed input text. Ultimately, a display of several videos playing in succession may be seen on the screen.

Implementation and outcomes

Step 1: Comprehension of business

The life cycle's beginning is the business need. As a result, it's crucial to know what the problem statement is and to ask the customer the correct questions so that we can fully comprehend the data and draw insightful conclusions from it.

Despite the fact that we have all the technology we need to live comfortably, the success of any project still depends on the calibre of the dataset queries that are posed.Every industry and domain operates under a set of guidelines and objectives. We should be able to comprehend the business in order to get the appropriate data. To narrow down the dataset and ensure proper data gathering, ask questions about it.

step 2: Data gathering

Finding the person who knows what data to collect and when to collect it based on the question to be answered is the first stage in the lifecycle of a data science project. Someone who understands the true distinctions between the many data sets that are accessible and can make tough judgements about an organization's data investment plan will be the best candidate for the position.

BIG DATA, the first word we used, applies here. Any data that is too large or complex to process is referred to as big data. Big data may not always imply vast scientific data. Four separate features define big data, and if your data demonstrates one of these characteristics, it is big enough to be referred to as such.

Data volume: Terabytes, Structured, semi-structured, and unstructured data are available in streams with high throughput. Veracity refers to the accuracy of the data being evaluated.

Step 3: Data pre-processing

also frequently referred to as the data wrangling stage. The most tedious and time-consuming activity involved in identifying various data quality issues, according to data scientists, is this one.We learn more about the data in this step and get it ready for more analysis. Is the data you gathered representative of the issue that needs to be solved? is a question addressed in the data comprehension phase of the data science approach.

You will inevitably finish this one task. Cleaning data basically entails eliminating errors from your data, such as blank fields and incorrect values, as well as establishing the proper data format and structuring data from raw files.Prepare the data as needed, removing any unnecessary columns and features. preparing data is is the longest and possibly most significant step in the entire life cycle. As good as your data are, so will be your model. Washing vegetables to get rid of surface pollutants is comparable to this. The time needed for data gathering, understanding, and preparation might range from 70% to 90% of the total project time.

At this point, you can also decide to return to the data collection phase if you believe the data is insufficient or unreliable to move on.

Step 4: Data analysis

VENTURE... VENTURE... VENTURE

Exploratory analysis is frequently referred to as a philosophy, and there are no set methods for conducting it. In terms of data exploration, there are no short cuts.Keep in mind that the calibre of your inputs determines the calibre of your output. Spending a lot of time and energy here makes sense once you have your business hypothesis prepared.Many people examine the data statistics, such as mean, median, etc., to better understand the data. Moreover, charts like the histogram, spectrum analysis, population distribution, etc. are used to examine the distribution of the data. We now plan to analyse the data using analytics. Depending on the circumstances, numerous forms of data analytics may be carried out on the data.

step 5:Data modelling

Almost all data scientists appear to find this stage to be the most fascinating. It is frequently referred to as "a stage where the magic happens." But keep in mind that magic only works with the right tools and methods. Data preparation is the method in data science, while "Data" is the prop. Thus, give the preceding stages enough time before moving on to this one.

To discover patterns or behaviours in data, modelling is utilised. We can benefit from these patterns in one of two ways:

1. Descriptive modelling (Unsupervised learning) – Recommender systems that predict whether a user will enjoy the movies Inception or Matrix based on their previous viewing preferences.

2. Predictive modelling (Supervised Learning) – Predictions are obtained by this method using linear regression, for example, we could forecast stock exchange prices for future trends.

Step 6: Model Assessment

Which dataset to use to measure the performance of the machine learning model is a frequent query among experts when assessing the performance of a machine learning model. Though useful, analysing performance metrics on the trained dataset is not necessarily accurate because the results could be too optimistic because the model has already been tuned to the training dataset. In order to determine the optimum model based on model accuracy and over-fitting, machine learning model performances should be assessed and compared using validation and test sets.

The choice of models could be made based on the business situation. It is critical to define the goal, including whether it is a clustering challenge, time series forecasting, regression or prediction problem, or classification problem.

The model might be used once the problem type has been determined.

Some illustrations of classification metrics:

1. Correct classification

2. Confusability matrix

3. Loss in Logarithms (Log Loss)

4. Curve's underside (AUC)

F-Measure 5. (F1 Score)

6. Precision

7. Recall

step 7:Model Deployment

Only having technical abilities is insufficient for this process. You must have the ability to communicate clearly and effectively, which is a necessary talent. If your audience doesn't respond to your presentation by taking some sort of action, your message was ineffective. It ought to be appropriate for inquiries about business. Both the organisation and the stakeholders should find value in it. The audience should be motivated to take action by the presentation through visualisation. Keep in mind that the audience you present to won't have any technical knowledge, so how you convey the content is crucial.

step 8:Action taking

Data science has the ability to perform predictive analytics and prescriptive analytics, as demonstrated by the model's actionable findings. This offers us the ability to discover ways to either duplicate positive results or avoid negative ones.We reach a point where we are able to make decisions about any current business problem based on all the insights we have gleaned through data observation or the output of a machine learning model.

Discussion

According to the grammar of Indian Sign Language, the verb and object positions in the given phrases must be switched, therefore the given input text is parsed using the StanfordParser to identify the subject, verb, and object in the given sentences. The system must then lemmatize each word because sign language does not follow tense, but lemmatization is dependent on a corpus, and we used the Wordnet corpus, which does not always stem all verb forms. As a result, it is necessary to edit the video and sigml files to address this issue.

Next, stopwords are looked for in the output generated and removed if any are found. The ISL statement is then put through two different processes to a sign language generator

We examined each of these techniques individually and developed the following analyses for each:

1. visual

Finding the appropriate videos and combining them into one cohesive video takes time. The absence of study in this area prevented the video output from being consistent. Every word may not have a word file counterpart, so such words are skipped.

2. Artificial Animation

Comparatively speaking, it takes less time to generate animation by mapping the sigma notations. This guarantees coherence because the same character acts out the entire statement. But unlike videos, these animations don't have the same sense of realism. For such, finger spelling is employed.

Conclusion and future scope:

The main goals of our paper are to highlight the value of ISL translation systems and to create a practical solution that will convert voice to ISL. Other widely used languages, such as American Sign Language, have been the subject of extensive research, although ISL has received less attention in this area. The creation of this system will aid in bridging the communication gap between the community of hearing and speech impaired people and the community of able-bodied people. The development of a society as a whole is facilitated by the inclusion of all communities.

In ISL, there are a maximum of 2700 dynamic motions. As many gestures as possible will be incorporated into our system over time. There must be an ongoing, protracted In order to make this product genuinely comprehensive and helpful, we regularly update our gesture library. Correctly identifying nouns, pronouns, tenses, prepositions, and other components of speech is the main issue in ISL to English 4 conversion. Better communication is required when identifying the sentence's tone and the emotions that go with it. A feedback mechanism will guarantee that our system keeps improving. When translating voice to Indian Sign Language, we aim for 100% accuracy. All around India, several different dialects are spoken. The fact that each state and province has its unique dialect makes this translation more challenging than before. Provisions for all the various languages are made by machine learning and a very strong, improved natural language processing dialects must be created. Native Language Support is essential for any product to be successful in India. There isn't a real, complete ISL database out there. We want to make it open source and add all of our ISL gesture data. For both the general public and other scholars and developers, this would be a really helpful resource.

Usually, persons who are dumb are unable to interact normally with the broader populace. The only options available to them are to use an interpreter or a visual medium of communication. Also, an interpreter cannot always be available, so we hope that our document will assist reduce the need for interpreters. The system must fully comprehend the context and tonality of input speech could be expanded to incorporate understanding of body language and facial expressions as well. A web application and a mobile application will broaden the audience. By making the technology open source, more and more people will be able to record signs for corresponding words and add them to the database already in place. In order to collect more sign language translations for various words, the data set will grow in size as a result.

References

[1] Indian Sign Language text generation from English/Hindi Text, Pawan Kumar, Savita Khatri et al. International Journal of Recent Research Aspects (IJRRA) ISSN: 2349-7688, Special Issue: Conscientious and Unimpeachable Technologies 2016, pp. 30-33

[2] National Sample Survey Organization. Disabled persons in India. NSS 58th round (July December 2002) Report no. 485 (58/ 26/ 1).New Delhi: National Sample Survey Organization, Ministry of Statistics and Program Implementation, Government of India, 2003.

[3] M.Suresh Anand, A.Kumaresan, Dr.N.Mohan Kumar, An Integrated Two Way ISL (Indian Sign Language) Translation System A New Approach, Volume 4, No. 2, Jan-Feb 2013 International Journal of Advanced Research in Computer Science.

[4] Hee-Deok Yang, Stan Sclaroff,and Seong-Whan Lee, Sign Language Spotting with a Threshold Model Based on Conditional Random Fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 7,pp.1264- 1277, July 2009.

[5] Archana S. Ghotkar, Rucha Khatal , Sanjana Khupase, Surbhi Asati Mithila Hadap, Hand Gesture Recognition for Indian Sign Language, 2012 International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 12, 2012, Coimbatore, INDIA.

[6] Yanan Xu, Yunhai Dai, Review of Hand Gesture Recognition Study and Application, Contemporary Engineering Sciences, Vol. 10, 2017, no. 8, 375 384.

[7] Ananya Choudhury, Anjan Kumar Talukdar, Kandarpa Kumar Sarma, A Conditional Random Field Based Indian Sign Language Recognition System under Complex Background,2014 Fourth International Conference on Communication Systems and Network Technologies.

[8] Gajanan K. Kharate, Archana S. Ghotkar, Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction.

[9] Dr.C. Chandrasekar, G.T. Shrivakshan, A Comparison of various Edge Detection Techniques used in Image Processing, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012.

[10] Samta Gupta , Susmita Ghosh Mazumdar, Sobel Edge Detection Algorithm, International Journal of Computer Science and Management Research.

[11] Lei Yang,Xiaoyu Wu, Dewei Zhao, Jun Zhai, Hui Li, An Improved Prewitt Algorithm for edge detection based on noised image, 2011 4th International Congress on Image and Signal Processing.

[12] Shilpa Bhople , Prof. R. R. Itkarkar , Pooja Bhoi, International Journal of Electrical Electronics Computer Science Engineering Volume 2, Issue 3 (June, 2015) — E-ISSN : 2348-2273 — P-ISSN : 2454-1222.

[13] Bag, S.; Sanyal, G., ‖An efficient face recognition approach using PCA and minimum distance classifier,‖ Image Information Processing (ICIIP), 2011 International Conference on , vol., no., pp.1,6, 3-5 Nov. 2011.

[14] G. R. S. Murthy R. S. Jadon, A review of vision based hand gestures recognition‖, International Journal of Information Technology and Knowledge Management July-December 2009, Volume 2, No. 2, pp. 405-

410.

[15] A. El-Sawah, N. Georganas, E. Petriu, A prototype for 3-D hand tracking and gesture estimation, IEEE Transactions on Instrumentation and Measurement, 57(8):16271636, 2008.