

Audio to Sign Language Translator Using Python

Mrs.B.Sivaranjani, M.Sc., M.Phil., (Ph.D.,)

Assistant Professor, Department of Information Technology

J.Sathya Priya, E.Ghanishka, V.Kamal Raj, Mohana Priya

UG Student, Department of Information Technology, Dr. N.G.P. Arts and Science College

ABSTRACT

This project is based on converting the audio signals receiver to text using speech to text API. Speech to text conversion comprises of small, medium and large vocabulary conversions. Such systems process or accept the voice which then gets converted to their respective text. This paper gives a comparative analysis of the technologies used in small, medium, and large vocabulary Speech Recognition System. The comparative study determines the benefits and liabilities of all the approaches so far. The experiment shows the role of language model in improving the accuracy of speech to text conversion system. We experiments the speech data with noisy sentences and incomplete words. The results show a prominent result for randomly chosen sentences compared to sequential set of sentences.

INTRODUCTION

It is easy to find a wide number of sign languages all over the world and almost every spoken language has its respective sign language, so there are about more than 200 languages available. 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).This section is not going to go further with details of a single sign language because each sign language has its own rules. The next section will aim to give a general description of the shared or common characteristics between the different sign languages: origin, phonology, and syntax.

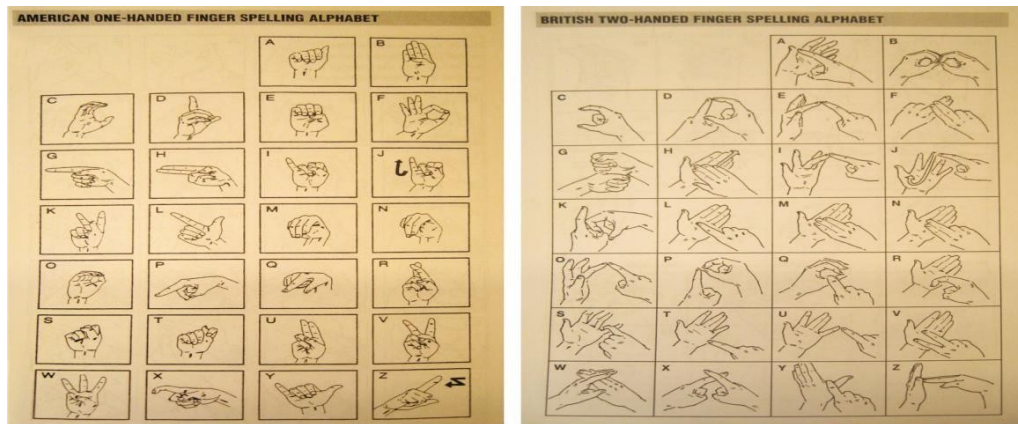


Fig1.1:- Block diagram of Sign Language system

Text to sign language conversion is mainly focused on communication between ordinary people and ordinary people and deaf-mute people. Sign language paves the way for deaf-mute people to communicate. Sign language is a visual language that is used by deaf and dumb as their mother tongue. It is figure out about 240 sign language have exist for spoken language in the world. Sign language is a type of language that uses hand movements, facial expressions and body language to communicate. It is used by the people who are deaf and people who can hear but cannot speak. The Conversion system consists of following steps:-

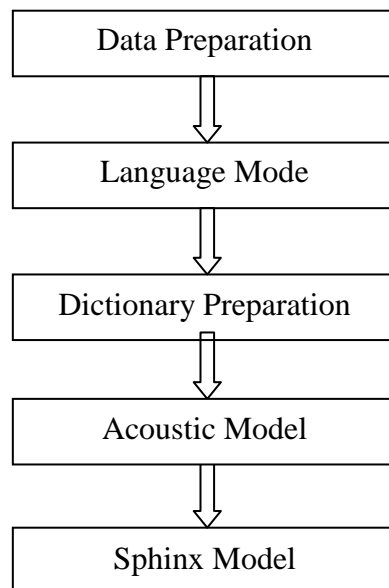


Fig1.2:- Block diagram of Conversion system

DATA PREPARATION

The corpus used for the system is the publicly available corpus. It contains a total of 1000 sentences about general information. The system is trained with 1000 sentences and tested 150 sentences.

LANGUAGE MODE

A Language model comprises of a large set of words together with its chances of occurrence. The model extracts the number of unigram bigram and trigrams of the corpus and calculates the probability of each unigram bigram and trigram. These statistical results are used by the decoder to predict the possible combination of words and phrases. It helps to achieve faster execution and higher accuracy by constraining the search in a decoder by limiting the number of possible words that need to be considered during the search.

DICTIONARY PREPARATION

Dictionary provides the data to map vocabulary words to sequence of phonemes to the system. Uses Letter-only phone names without special symbols which simplifies the system. Dictionary should contain all the words needed to be recognized by the recognizer.

ACOUSTIC MODEL

Acoustic model is a file which contains statistical representation of each of individual sounds that make up a word. An acoustic model is created from a speech corpus using training algorithms. In Sphinx it is done using Sphinx train module. This part gives the output in the form of a configuration file. The parameters written in configuration file are used by the decoder to generate the acoustic model for a given language.

SPHINX TRAIN (Open source toolkit for speech recognition)

Training is performed when there is need to create an Acoustic model for a new language. Knowledge on the phonetic structure of the language should be there to perform the training. Once the training is done it creates the database and by running the sphinx train the speech recognition files can be created.

TRAINING ALGORITHM

Acoustic model is a file which contains statistical representation of each of individual sounds that make up a word. An acoustic model is created from a speech corpus using

training algorithms. In Sphinx it is done using Sphinx train module. This part gives the output in the form of a configuration file. The parameters written in configuration file are used by the decoder to generate the acoustic model for a given language.

EXISTING SYSTEM

This approach should be capable to recognize the speech and convert the input audio into text. Likewise, this problem related to several problems. Speech recognition is an interesting application of digital signal processing which has real world applications. This method is also used in automation of many tasks which previously needed the human interaction, like identifying spoken commands to perform things like closing a door or switching on lights.

DE-METRIC

Complex speech patterns can be recognized as well. For instance, there are quite a few appropriate speech recognitions which can actually take up speech at decent speed and later convert it to the text format and hence no typing would be required to generate a document. Even after such successful software landing in the market however, current efforts are not yet meeting the 100% human speech recognition.

PROPOSED METHODOLOGY

Tensor layer was replaced with single sigmoid hidden layer by Hutchinson, Deng and Yu in the stacking networks. The performance was worst when the configuration in which only the bottom (first) layer was replaced with the DP layer. The performance was best and achieved more than 1% absolute reduction over the DNN when the configurations replaced the top hidden layer with the DP layer performs. This concludes the DP layers are suited to perform on binary features, consistent in findings from.



METRICS

On a voice search task and the Switchboard (SWB) phone-call transcription task it is found that CD-DNN-HMMs have achieved 16% and 33% relative recognition error reduction over strong, discriminatively trained CD- GMM-HMMs.

CONCLUSION

Sign language is one of the useful tools to ease the communication between the deaf and mute communities and normal society. Though sign language can be implemented to communicate, the target person must have an idea of the sign language which is not possible always. This was meant to be a prototype to check the feasibility of recognizing sign language. The normal people can communicate with deaf or dumb using sign language and the text will be converted to sign images.

REFERENCES

1. Amit kumar shinde and Ramesh Khagalkar “sign language to text and vice versa recognition using computer vision in Marathi” International journal of computer Application (0975-8887) National conference on advanced on computing (NCAC 2015).
2. Sulabha M Naik Mahendra S Naik Akriti Sharma " Rehabilitation of hearing impaired children in India"International Journal of Advanced Research in Computer and Communication Engineering.
3. Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak "Study of Sign Language Translation using Gesture Recognition" International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.
4. Christopher A.N. Kurz "The pedagogical struggle of mathematics education for the deaf during the late nineteen century: Mental Arithmetic and conceptual understanding" Rochester Institute of Technology, Rochester, NY USA. Interactive Educational Multimedia, Number 10 (April 2005), pp. 54-65.
5. Foez M. Rahim, Tamnun E Mursalin, Nasrin Sultana “Intelligent Sign Language Verification System Using Image Processing, clustering and Neural Network Concepts” American International University of Liberal Arts-Bangladesh.
6. Shweta Doura, Dr . M.M.Sharmab "the Recognition of Alphabets of Indian Sign Language by Sugeno type Fuzzy Neural Network"International Journal of Scientific

Engineering and Technology (ISSN : 2277-1581) Volume 2 Issue 5, pp : 336-341 1
May 2013.

7. Neha V. Tavari A. V. Deorankar Dr. P. N. Chatur" A Review of Literature on Hand Gesture Recognition for Indian Sign Language"International Journal of Advance Research in Computer Science and Management Studies Volume 1, Issue 7, December 2013.
8. Vajjarapu Lavanya, Akulapraavin, M.S., Madhan Mohan" Hand Gesture Recognition And Voice Conversion System Using Sign Language Transcription System" ISSN : 2230-7109 (Online) | ISSN : 2230-9543 (Print) IJECT Vol. 5, Issue 4, Oct - Dec 2014.
9. Sanna K., Juha K., Jani M. and Johan M (2006), Visualization of Hand Gestures for Pervasive Computing Environments, in the Proceedings of the working conference on advanced visual interfaces, ACM, Italy, p. 480-483.
10. Jani M., Juha K., Panu K., and Sanna K. (2004). Enabling fast and effortless customization in accelerometer based gesture interaction, in the Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia. ACM, Finland. P. 25-31.
11. Divyanshee Mertiya, Ayush Dadhich, Bhaskar Verma, DipeshPatidar “A Speaking module for Deaf and Dumb”, student, assistant professor Department of Electronics & comm. Poornima Institute of Engineering and Technology, Jaipur, Rajasthan, India.
12. T. Kapuscinski and M. Wysocki, “Hand Gesture Recognition for Man-Machine interaction”, Second Workshop on Robot Motion and Control, October 18-20, 2001, pp. 91-96.
13. D. Y. Huang, W. C. Hu, and S. H. Chang, “Vision-based Hand Gesture Recognition Using PCA+Gabor Filters and SVM”, IEEE Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009, pp. 1-4.
14. C. Yu, X. Wang, H. Huang, J. Shen, and K. Wu, “Vision-Based Hand Gesture Recognition Using Combinational Features”, IEEE Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 543-546.