Stuttered Speech Recognition

Anshuman S. Dhamoon, Sarthak Khanna, Sankar Ganesh S | SENSE

Motivation/ Introduction

Stuttering is one of the major speech impediments that people are suffering from, in today's world. It is a disorder which affects the fluency of speech by involuntary repetitions or prolongations of words, syllables, etc or involuntary silent intervals. Of all these, repetition is the most common and prominent characteristic of stuttering. In today's world, about 1% of the total population (or 70 million people) stutter, which shows that it is a major problem.

SCOPE of the Project

A few institutions have researched on detection of stuttering using acoustic level analysis methods. A better method than the above is to create a stuttering detection system based on a large vocabulary of continuous speech recognition, but a technique like this is very resource and power intensive, and hence, cannot be used practically.

We observed, that when a person stutters, there is a decrease in the amplitude in the person's voice signal and we used this, to eliminate the repetitions, elongations and silent intervals to produce a better speech recognition system.

Methodology

A simple solution to overcome the enormous computation power and gigantic datasets, is to use amplitude as a factor, to remove stuttering. We observed that when a person stutters, the amplitude of his speech drops significantly, when compared to the rest, and hence, could be used to filter out the stutter from the sample.

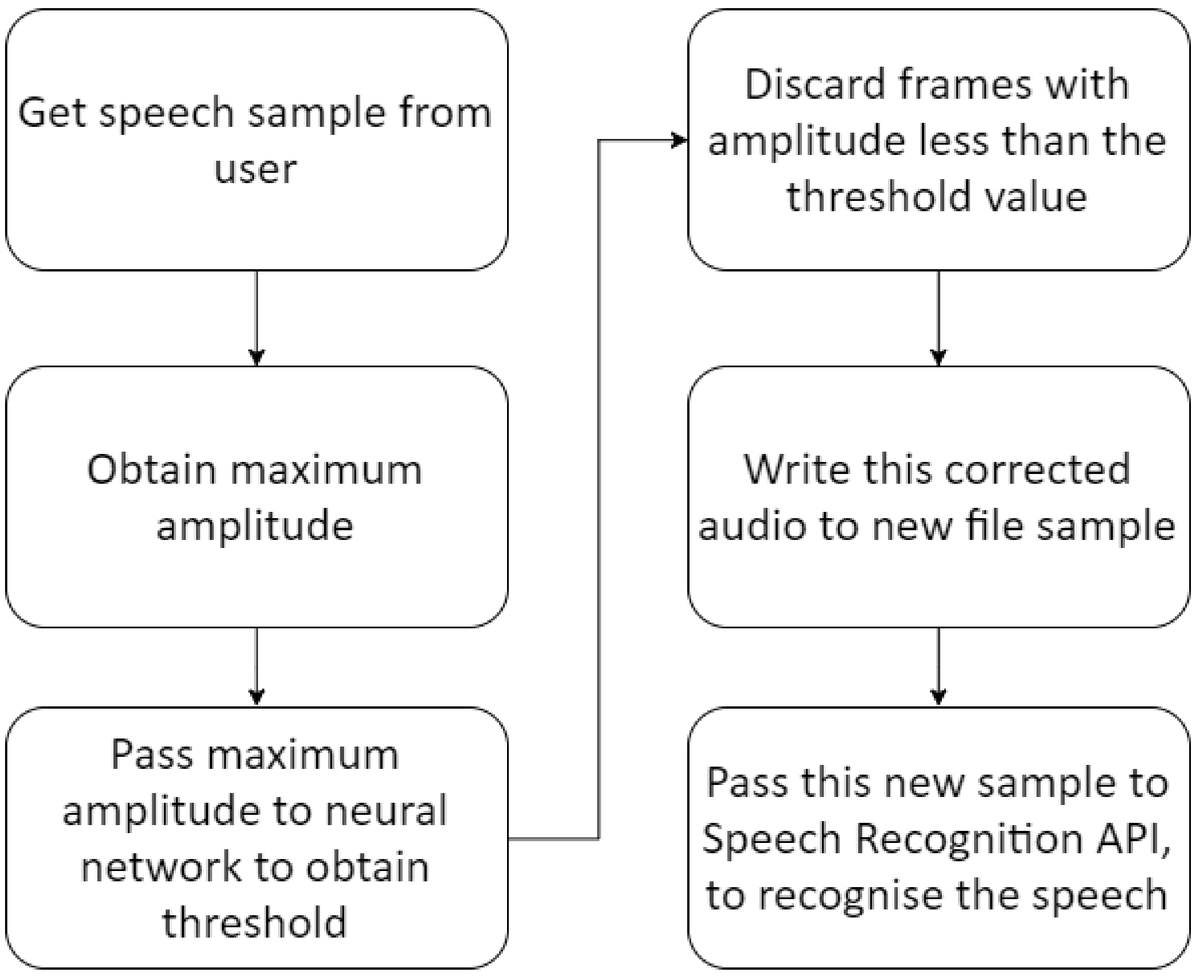
So, we checked a lot of stuttered speech datasets for the above, and noticed that the threshold needed for removing the stutter from the sample largely depended on the maximum amplitude observed in the sample.

We created a neural network that would only take the maximum amplitude of the sample as an input and give the required threshold as an output. We initially selected the thresholds manually, by looking at the plots of the speech samples and after acquiring a large enough dataset, used those values to train a neural network using back propagation algorithm. After the training was complete, the threshold output given by the network was accurate enough for us to clean the stuttered sample enough, that it could be easily recognized.

Since we eliminated the stuttering, we now simply had to pass a normal speech to an already existing system. Google is the leading company when it comes to speech recognition and they have spent years, perfecting it for different types of words, sequences, accents, etc. And the system that they have created has been open sourced for the use of general public.

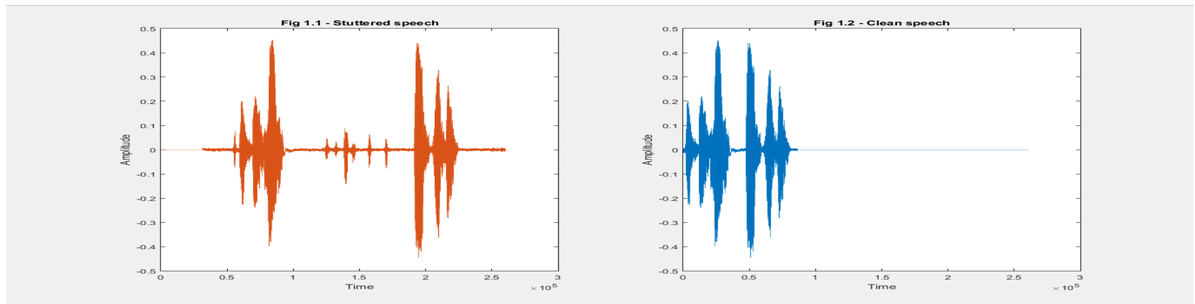
So, after acquiring the clean speech, we created a python script to which we could pass this speech, from MATLAB. Then we call the services of Google's system, pass it our speech and the system does the rest of the work by analyzing the speech and returning the result.

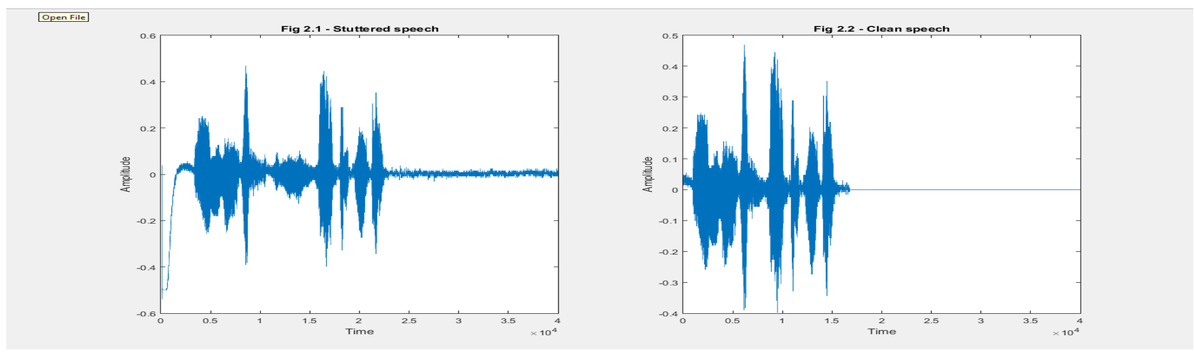
The following flowchart explains better.

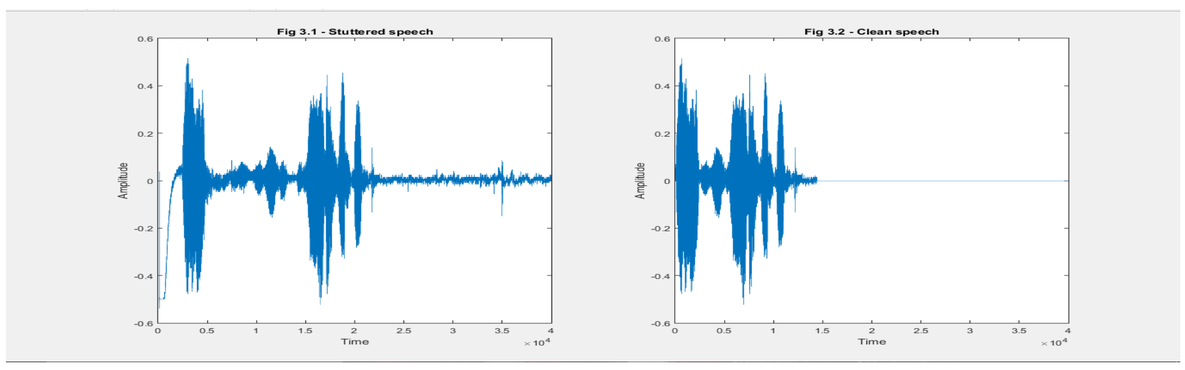


Results

The following plots show the stuttered speech and speech after the stutter has been removed for three different samples.







From the above plots, we can observe the difference in the amplitudes when stuttering occurs and how we used threshold values, which we obtained from the neural network that we trained, are used to clean the sample.

We tested using a total of 50 test cases, with 42 of them being detected correctly. A summary of the result obtained is shown in the table below.

Total # of test samples	Predicted correctly	Predicted incorrectly
50	42	8

K = P/n

Where,

K - Accuracy

P - Correctly predicted test samples

n - Total samples

Therefore, putting values form Table 1 into the equation we get an accuracy of:

Accuracy = 84.0 %

Conclusion/ Summary

The increasing usage of speech recognition systems by people has led to the ease of access in their day to day lives. However people with speech impairments like stuttering cannot benefit from these services. In order to make these above mentioned applications more universal we worked on a project that would actually help in solving a real world problem and we were successful in getting favorable outcomes.

Our stuttered speech correcting program was able to achieve an accuracy of 84% on 50 test samples. This shows promise for our approach and the accuracy can further be improved by increasing the number of training samples that we used in the training set.

Contact Details

anshumandhamoon@gmail.com, khannasarthak.1994@gmail.com

References

1. Overview of Automatic Stuttering Recognition System, Lim Sin Chee, Ooi Chia Ai, Sazali Yaacob.
2. Automatic Detection of Syllable Repetition in Read Speech for Objective Assessment of Stuttered Disfluencies, K. M. Ravikumar, Balakrishna Reddy, R. Rajagopal, and H. C. Nagaraj
3. Structure Analysis of Speech Signals and Filtering them from Noise, Anastasiya Andreevna Kuznetsova.