

DSP Final Report

VOP Detection using Spectral Peaks

Siddharth Gaur - 20171198

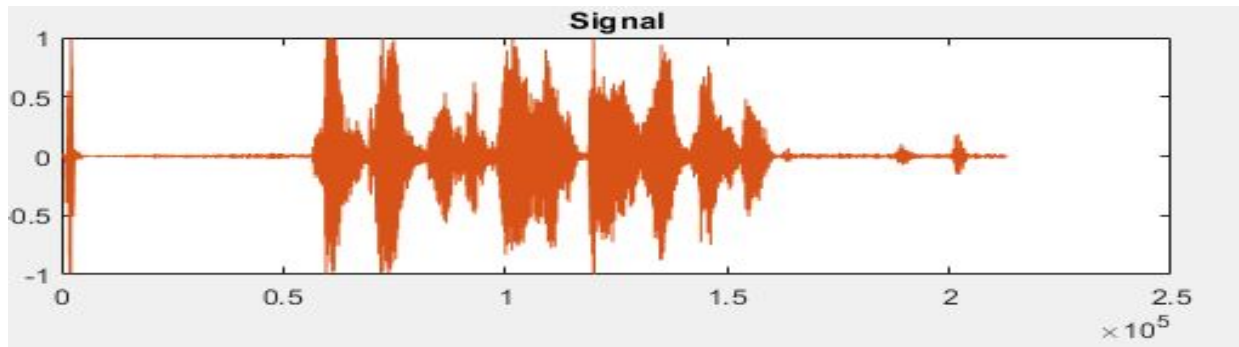
Mentor: Dr. Anil Kumar Vuppala

Project Description

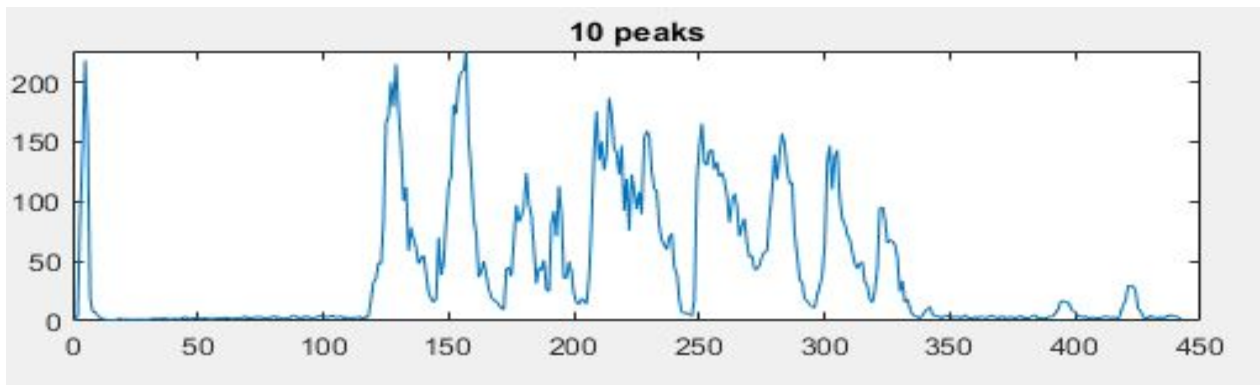
- Speech is produced as a sequence of sound units. These soundunits are produced as a result of changes in the vocal tract shape. Distinct vocal tract shapes are associated with the production of Vowel.
- Vowel onset point is the instant at which the beginning of vowel takes place during speech production.
- Speech signal is composed of voiced signal and unvoiced signal. Vowles along with few consonants are considered to be as voiced signal while others are unvoiced.
- Voiced signal tends to be louder like the vowels on the other hand, unvoiced signal tends to be more abrupt like the stop consonants.
- So whenever vowel is produced there must be a significant change in spectral energy.

Mathematical model and Results

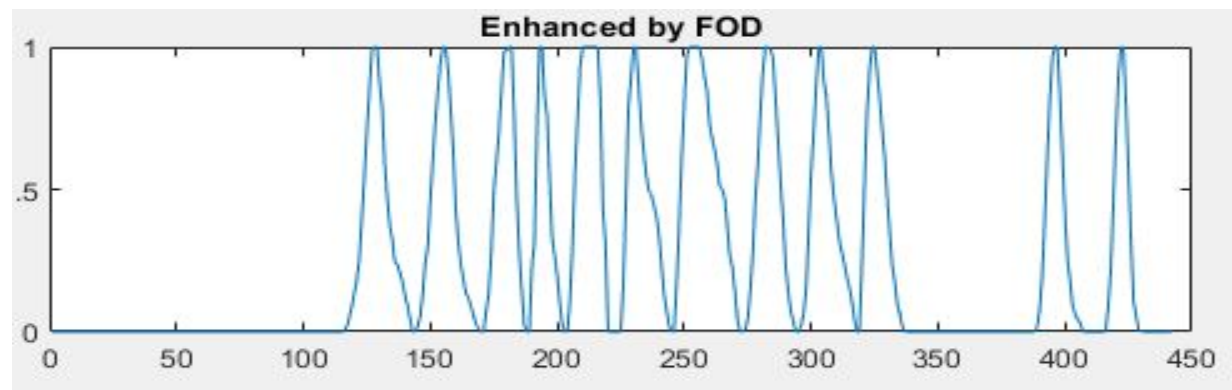
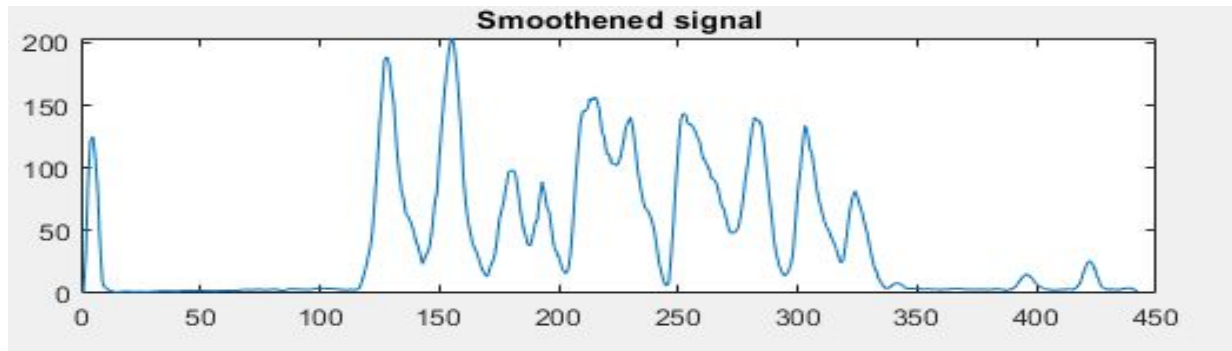
A Voice signal is taken as input : *“Don’t ask me to carry an oily rag like that”*



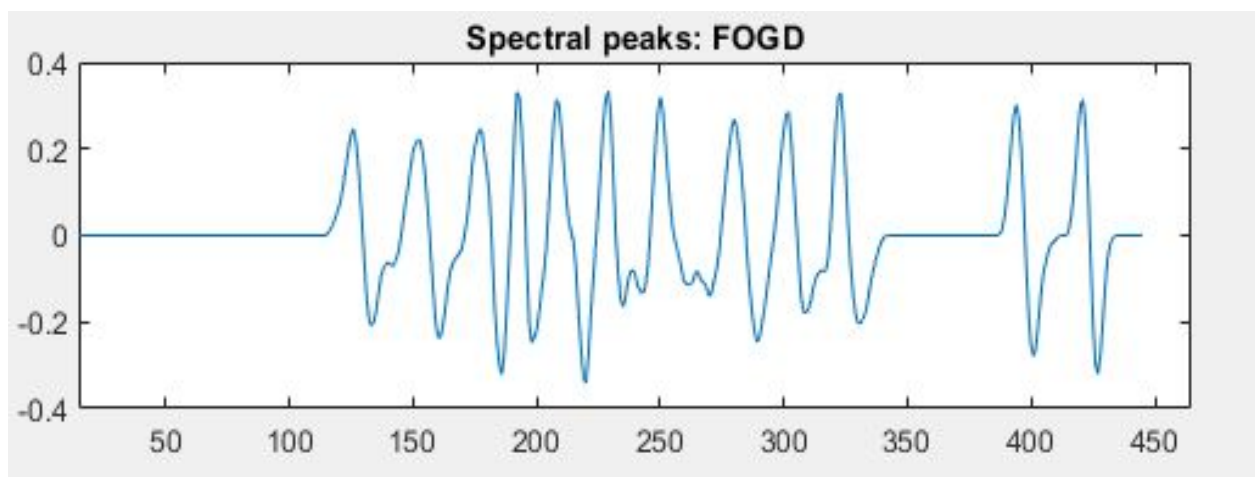
The speech signal is processed in a block of 20ms with a shift of 10ms. For each block of 20ms , 256 point DFT is computed and ten largest peaks from first 128 points are selected. The sum of these amplitudes is plotted as a function of time which represents the energy of spectral peaks.



The onset of vowel is observed as significant change in sum of ten peaks in the DFT spectrum. Signal is then smoothed and then enhanced using FOD (First order differentiator)



Further, enhancing is performed by convolving it with FOGD operator and the obtained output is the VOP evidence Plot using Spectral Peaks.



Conclusion

- Evidence plot is similar to that in the reference paper.
- Number of VOPs is the same as calculated manually (12).

Reference

- Vowel Onset Point Detection Using Source, Spectral Peaks, and Modulation Spectrum Energies.
(<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4802173>)