DSP PROJECT

VOWEL ONSET POINT DETECTION USING MODULATION SPECTRUM ENERGY

Mentor: Dr Anil Kumar Vuppala & Krishna Gurugubelli sir
Team Members:

Salay Jain (20171078)

Aradhya Tongia (20171049)

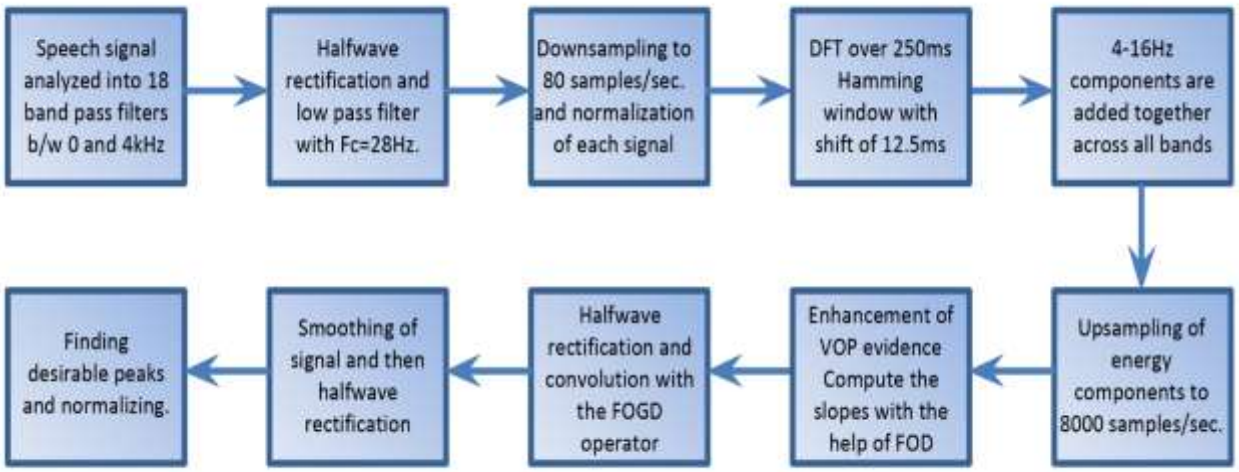
Vowel onset point (VOP):

- Vowel Onset Point is the instant at which the beginning of a vowel takes place during speech production. A speech signal is composed of voiced signal, unvoiced signal and noise. Vowels along with few consonants a considered as voiced while others as unvoiced.
- Voiced speech is produced by vibrations in vocal cords thus can have high energy compared to unvoiced speech which are low energy. There are significant changes occurring in the energies of excitation source, spectral peaks, and modulation spectrum at the point when a vowel is pronounced in speech. So our project is to detect Vowel onset point using modulation spectrum energy.

Significance of using modulation spectrum energy:

- Modulation components refer to the slowly varying temporal envelope components in speech. Temporal envelope of speech is dominated of low frequency components by several Hertz.
- Thus, the onset of vowel may be manifested as a significant change in the modulation spectrum energy level in the 4–16 Hz band.

Block diagram:



Results:

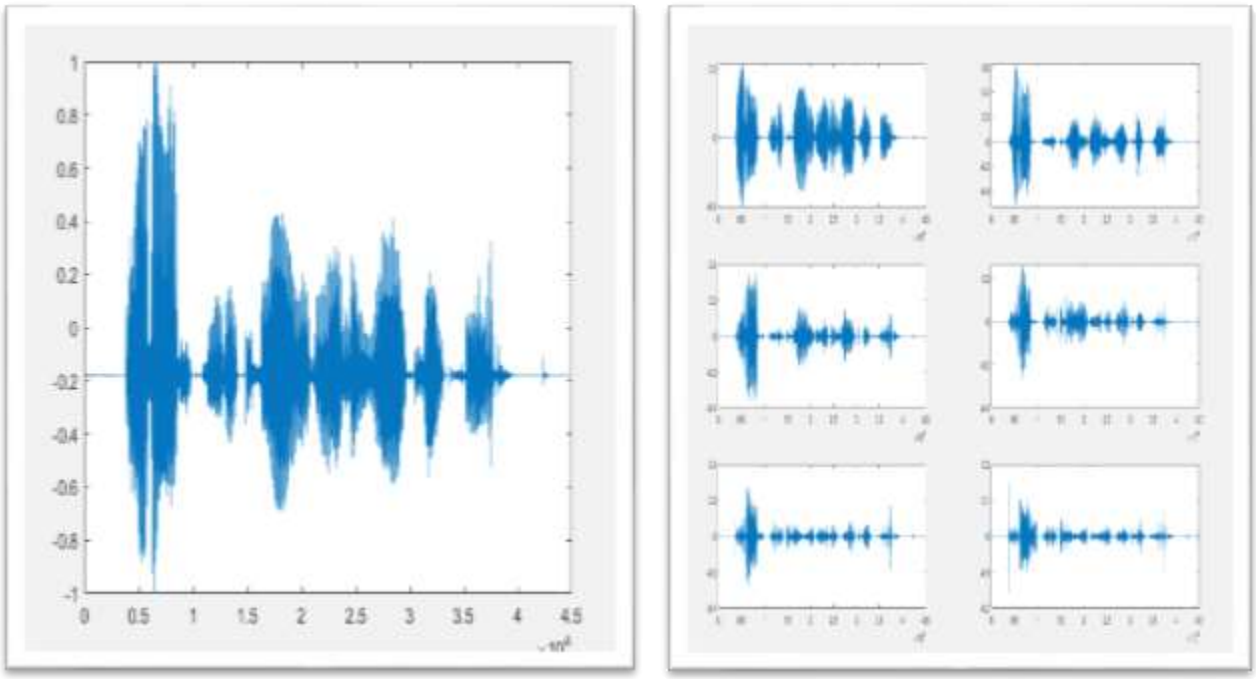


Figure 1:Speech signal “Dont_ask_me_to_carry_an_oily_rag_like_that”

Figure 2: Speech signal analysed to 18 bandpass filters between 0-4KHz.

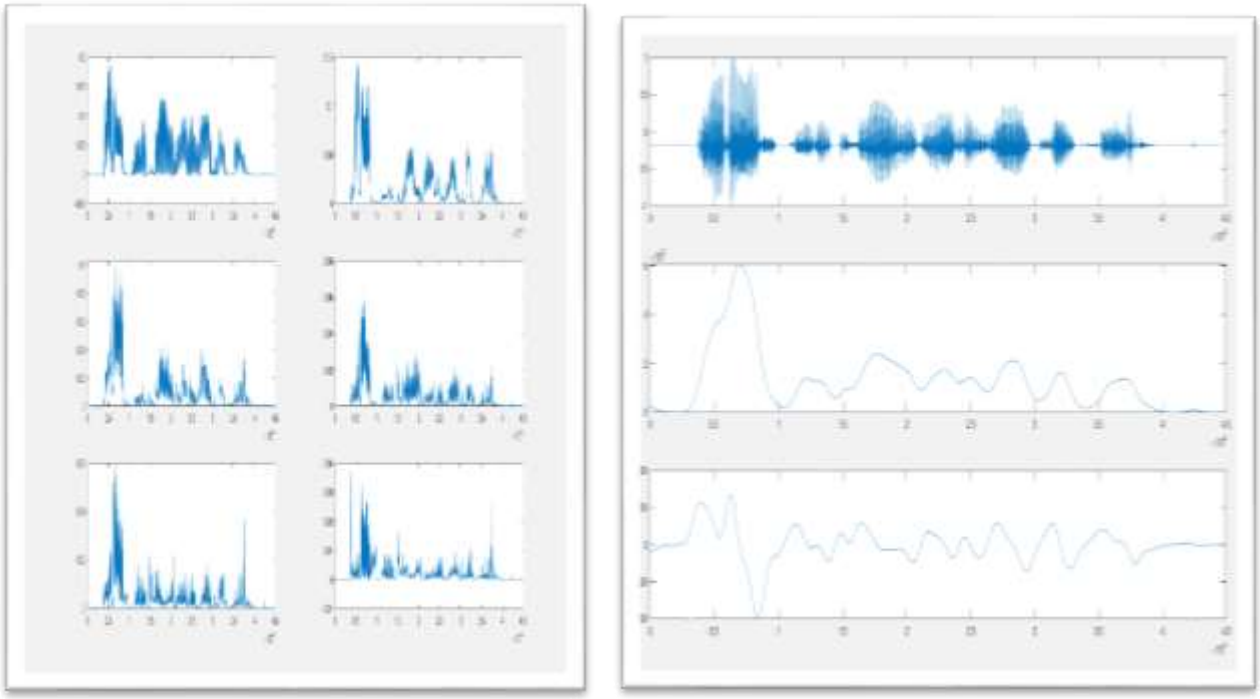


Figure 3: Half-wave rectification and low pass filter of 28Hz for all the 18 bands.

Figure 4: A)Speech signal to identify VOP’s.

B) Energy components of 4-16 Hz are added and smoothen for all the 18 bands.

C)First order Differential of the energy band after smoothening.

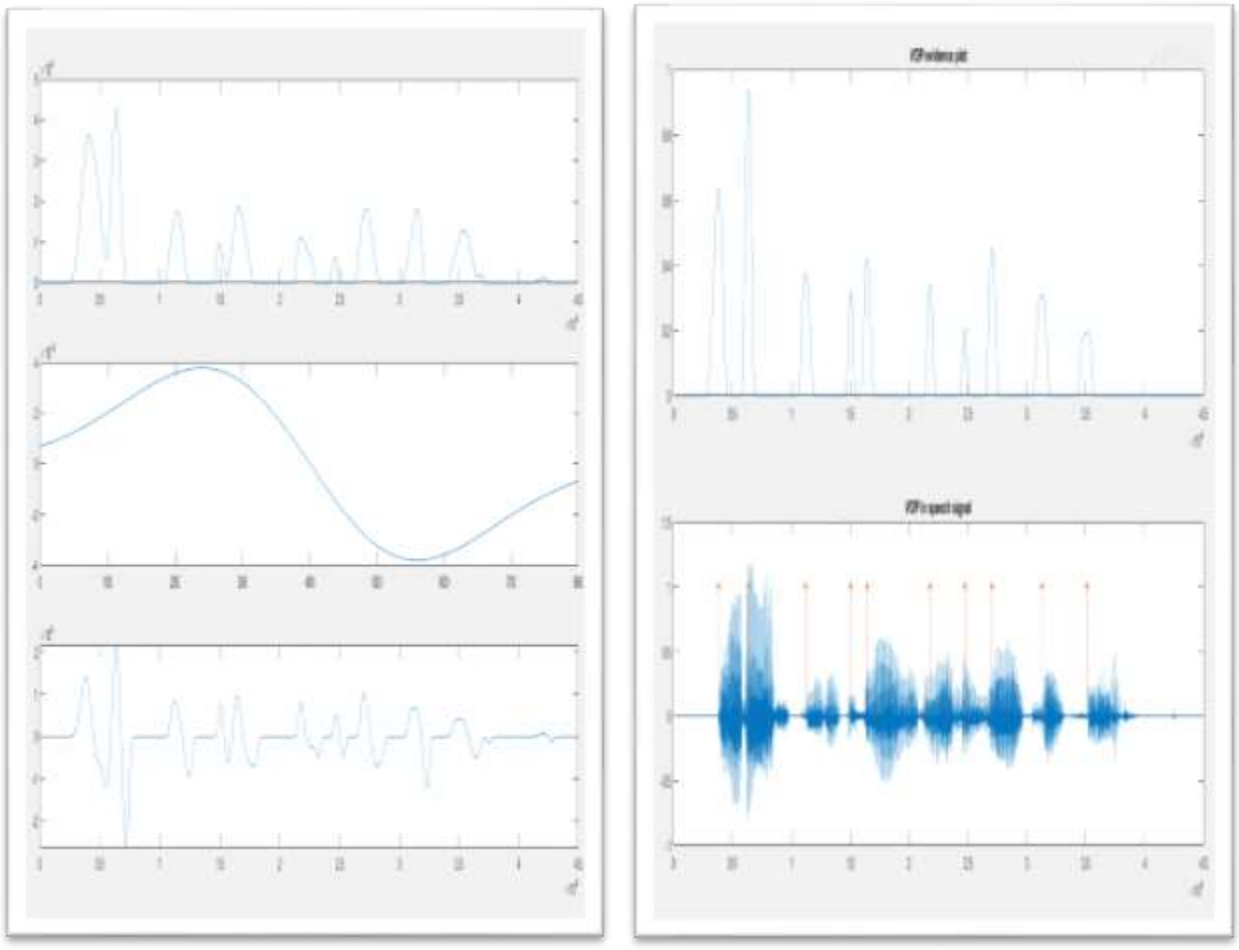


Figure 5: A) Sum of peaks for 10 mS window period is taken and rectified.
B) FOGD operator with which we will convolve the above signal to get accurate VOP's.
C) After applying first order gaussian differential operator we get enhanced VOP.

Figure 6: A) VOP evidence plot after half wave rectification and normalizing.
B) Peaks in VOP evidence plot gives us vowel onset point in our speech signal.Thus all orange '*' denotes the VOP in our signal.

- Thus we can see that we successfully recognized the VOP from our speech signal and missed only one of the VOP.
- We can see that if large noise is added than it will also be detected as VOP and thus this is the only defect in this process.
- Thus we are able to find the VOP in our speech signal successfully.
- This can be used for consonant vowel transitions, detection of end-points of speech utterance, also used as anchor points to extract features for speech recognition.
