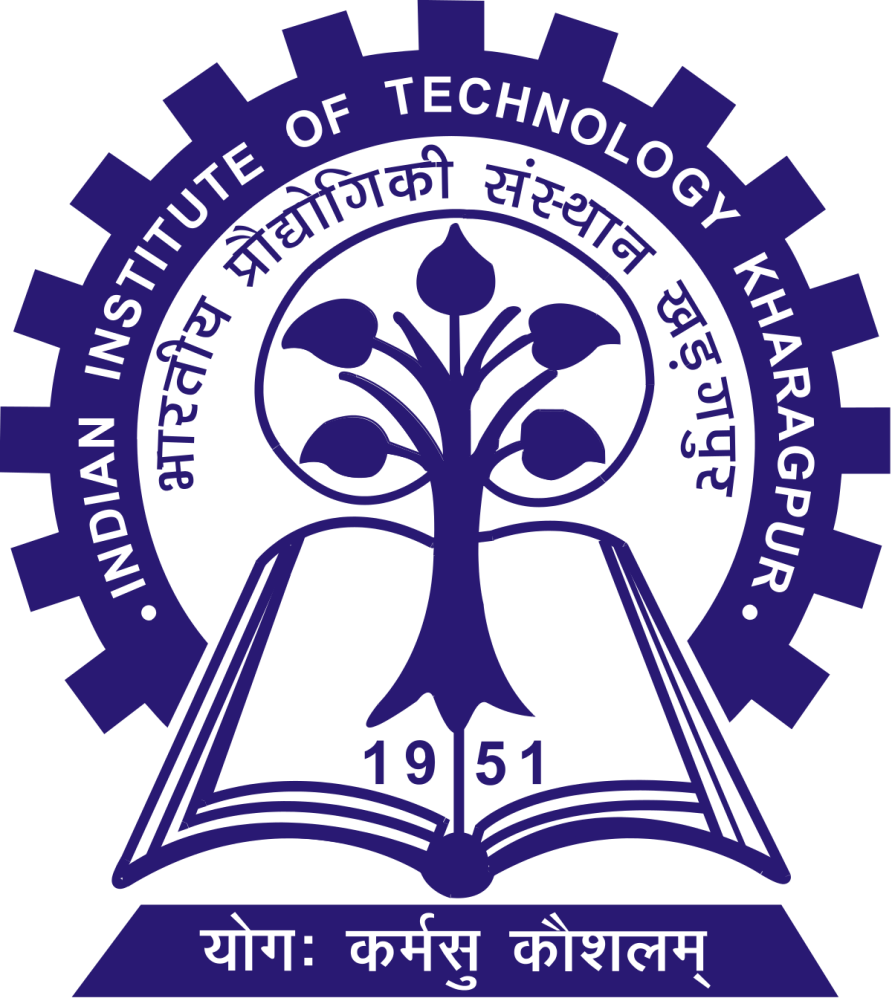
INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR



Department of Electronics & Electrical Communication

Engineering

M.Tech. First Year

Vision and Intelligent Systems

(VIS)

EC60064-Biomedical System Engineering and Automation

Assignment 5

Submitted by

Suraj Kumar

Roll No. = 22EC65R14

**DISCUSSION:**

1. Voiced and unvoiced signals can be segmented by thresholding the RMS value, turns count, and zero crossing rate.
2. So, using those three, we can have separation between the voiced and the unvoiced signal. However, in case of the voiced sound ‘E’ and ‘I’, when we look at the spectrum, we get that at certain frequencies, have the peaks.
3. For both of them in time domain, we get some repeated waveform and gives rise to the concentration of energy at certain frequency in the PSD. On the other hand, for the random kind of time domain waveform we get for the unvoiced signal. And, there is nothing specific in the time domain. In the same way, in the frequency domain, we do not get any peak in the PSD.
4. The RMS value of the voiced portion is more as compared to the unvoiced portion of the consonant part,
5. In terms of turns count and the zero crossing, both are higher for the unvoiced portion as compared to the voiced portion. These can help us to separate which portion is the voiced and which portion is unvoiced.
6. The derivative of the original signal, the number of zero-crossings increase for voiced, unvoiced, and silent portion. When we take the derivative signal, the zero crossing is increasing, it could be that because the DC bias which may not be visible in the plot, but small DC biases are there which are removed and that helps to get more zero crossings, and everywhere we get an increase and that makes the thing more uniform, but which is not a good thing. We wanted to have them separated so that we can use that zero-crossing rate to separate the 3 things out.
7. The RMS value, turns count, and the number of zero crossings are least for the silent portion as compared to the both the voiced and the unvoiced portion of the signal. So, what we get primarily that the turns count and the zero-crossing rate are similar. RMS value, turns count, and zero crossing, are less compared to the voiced and unvoiced portion of this signal.
8. Once we could take out the portion which is the speech portion using the RMS value and turns count, we can separate out the voice portion and the unvoiced portion. The voice portion will have high RMS and low turns count and the zero crossing whereas, the unvoiced portion will have high turns count and zero crossing and low RMS value.