Tim-domain analyses

03 September 2019 08:59 PM

Short-time energy
Short-time ZCR
Short-time Autocorrelation Function

Pitch detection challenges:

- Nonstationarity of speech, esp. at phone boundaries where formants change and waveform is modified drastically
- Irregularity in vocal fold vibration (jitter, shimmer)
- The wide range of possible F0 values (deciding PDA parameters such as window length)
- Interaction of F0 with vocal tract shape => formant structure interferes with F0 detection
- Degraded speech in noisy and reverberant environments-> periodicity is less evident

Pitch detection with the short-time spectrum: Harmonic Product Spectrum (HPS) method

$$P(n,\omega) = \prod_{r=1}^{k} |X(n,r\omega)|^2$$

$$\hat{P}(n,\omega) = 2\sum_{r=1}^{k} \log(\,|X(n,r\omega)|)$$