INDIAN INSTITUTE OF INFORMATION TECHNOLOGY

Dharwad

ज्ञानेन विकासः

Speech Enhancement

-Group:9

Kummara Bhargavi (18BEC024) and Kukkala Bharath (18BEC023)

Problem statement:

To remove the additive noise in the speech using **adaptive wiener filtering** technique.

- We have taken an input signal of N samples as reference signal.
- And added noise of same sample length to the input signal
- The problem is to extract the input signal from output signal (noisy signal) by eliminating the noise

Glimpse of adaptive wiener filtering:

Given the degraded speech signal, divide it into small segments. Find the filter transfer function for each segment and enhance each segment of the degraded speech. Finally, all the modified segments are combined to get enhanced speech.

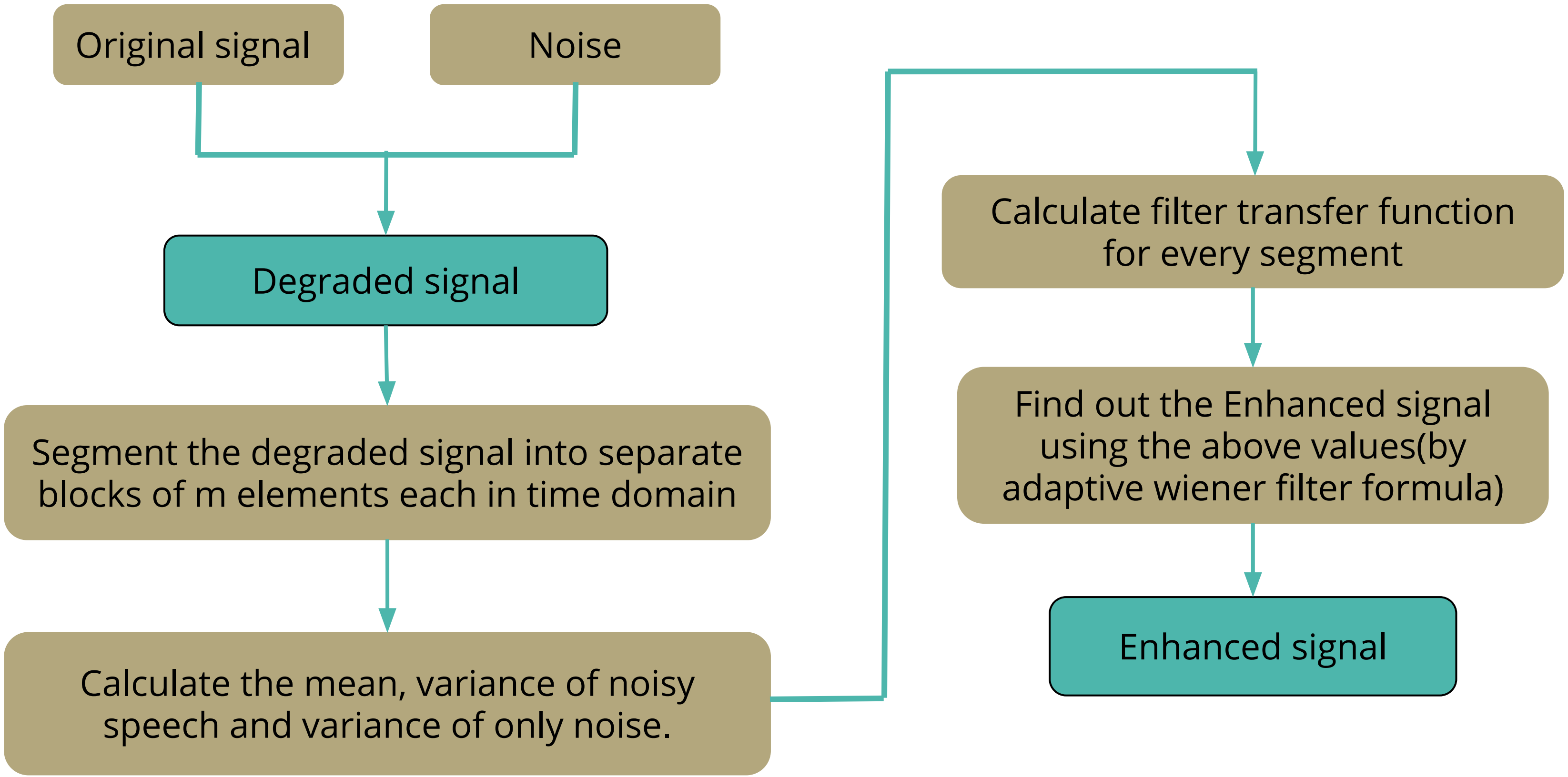
The designed adaptive wiener filter depends on the adaptation of the filter transfer function from sample to sample based on the speech signal statistics like mean and variance.

Application of adaptive filter:

Noise removal from the pilot's microphone in the airplane:

Due to high environmental noise produced by the airplane engines, the pilot's voice in the microphone is distorted with a huge amount of noise, and it can be very difficult to understand. Adaptive filters can be very helpful here to remove the noise.

Flowchart of implementation:



Technical details:

After adding noise to the clear speech signal, calculate estimates of mean and variance of degraded speech signal segment wise using the formulas below:

$$\widehat{m_s}(n) = \frac{1}{(2M+1)} \sum_{k=n-M}^{n+M} x(k)$$

$$\widehat{\sigma_x^2}(n) = \frac{1}{(2M+1)} \sum_{k=n-M}^{n+M} (x(k) - \widehat{m_x}(n))^2$$

Within every small segment, wiener filter transfer function is approximated by:

$$h(n) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2} \delta(n)$$

Continued...

The enhanced speech signal $\hat{s}(n)$ in this local segment can be expressed as:

$$\hat{s}(n) = m_x + (x(n) - m_x) * h(n)$$

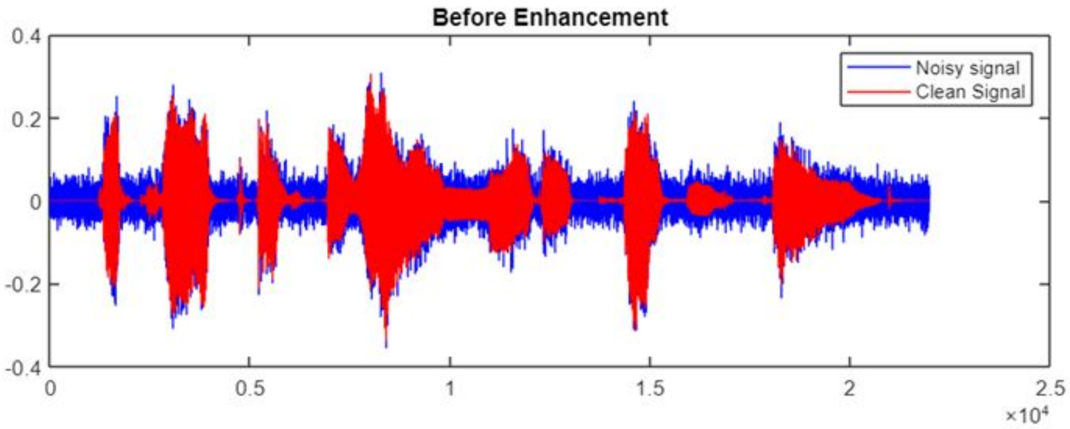
$$\hat{s}(n) = \widehat{m_s}(n) + \frac{\widehat{\sigma_s^2}(n)}{\widehat{\sigma_s^2}(n) + \widehat{\sigma_v^2}(n)}(x(n) - \widehat{m_s}(n))$$

where $\widehat{\sigma_s^2}(n) = \widehat{\sigma_x^2}(n) - \widehat{\sigma_v^2}(n) \quad when\ \widehat{\sigma_x^2}(n) > \widehat{\sigma_v^2}(n);\ otherwise = 0$

OUTPUT

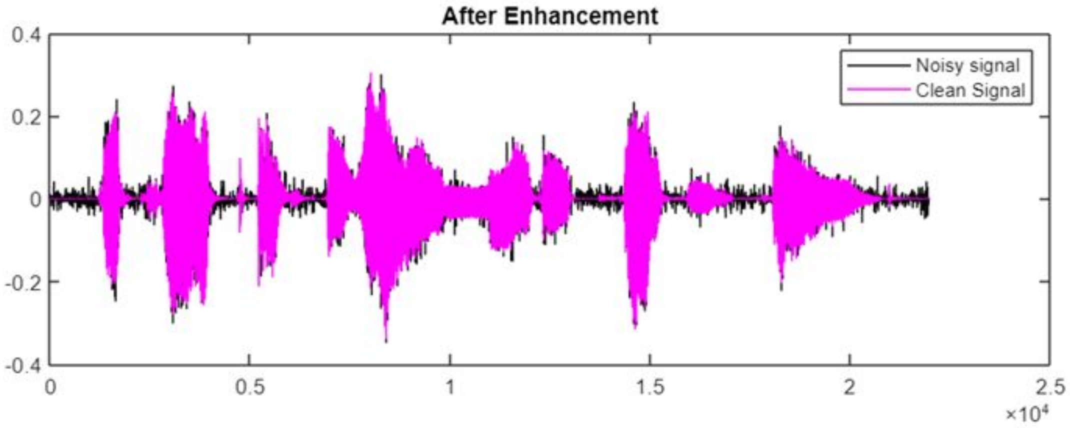
SNR before

Enhancement: 5dB



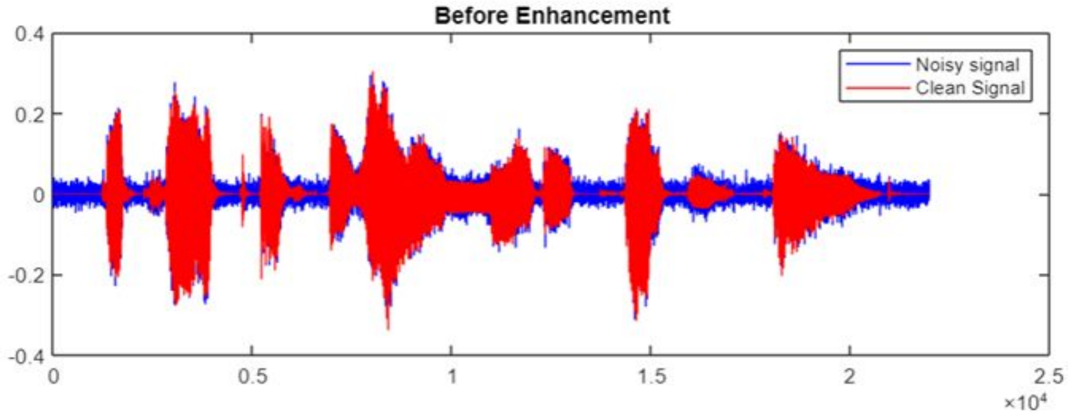
SNR after

Enhancement: 5.3158dB



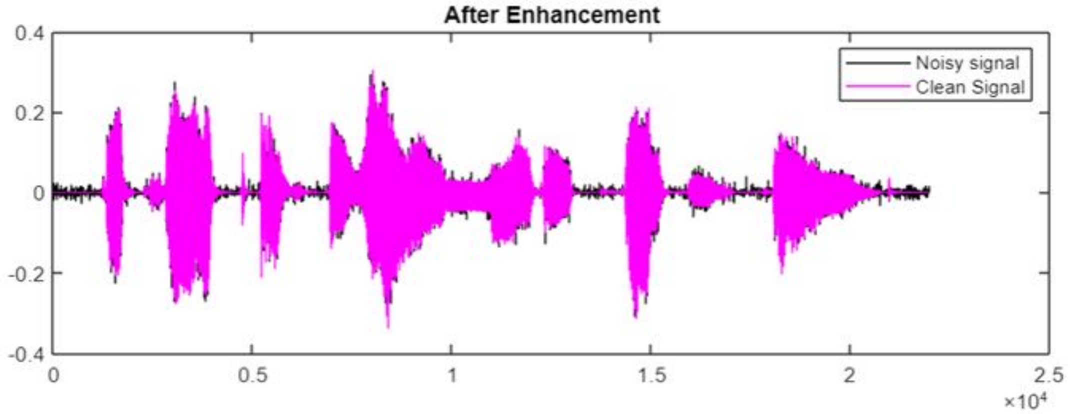
SNR before

Enhancement: 10 dB



SNR after

Enhancement: 10.4837 dB



SNR before

Enhancement: -5 dB

SNR after

Enhancement: -3.7943 dB