

DEEE725 Speech Signal Processing Lab

I. 배경잡음 제거하기

녹음환경에서는 추가되는 다양한 잡음들의 종류

① 주변환경 잡음: 음악, 자동차, 사람 등

④ white noise
(백색잡음)

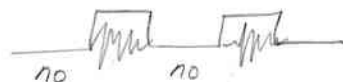
② AC noise: 50~60 Hz의 sinusoidal 잡음

③ dc noise: base level이 올라감.



가정

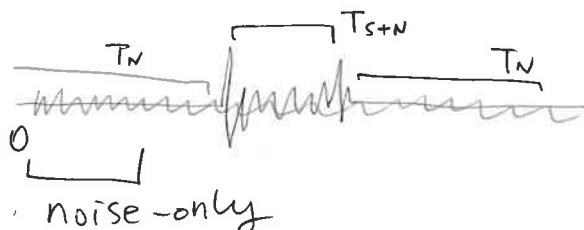
① speech signal은 시간에 따라 주파수 특성이 바뀌며 (non-stationary)
항상 존재하지 않는다.



② noise signal은 주파수 특성이 일정하다 (stationary)

③ noise-only 구간이 있다

가장 쉬운 방법으로 시작 시점으로부터 일정 시각까지는
음성이 없다고 가정한다. 예를 들어 0~0.5초는 noise-only



그러면 noise-only 구간에서 noise spectrum을 얻을 수 있다.

관측신호 $x[t]$ 은

$$x[t] = s[t] + n[t] \Leftrightarrow X(\omega) = S(\omega) + N(\omega)$$

$$\tilde{N}(\omega) = E[F(n[t])] \approx \frac{1}{T} \sum_{t=0}^{0.5} F(x[t])$$

power spectral density로 근사한다

$$\tilde{P}_{nn}(\omega) = |\tilde{N}(\omega)|^2 = \frac{1}{0.5} \sum_{t=0}^{0.5} x(\omega) \cdot \bar{x}(\omega)$$

complex conjugate of $x(\omega)$

Spectral Subtraction

$$|\tilde{S}(\omega)|^2 = |S(\omega) + N(\omega)|^2 - |\tilde{N}(\omega)|^2 = X(\omega) \cdot \bar{X}(\omega) - |\tilde{N}(\omega)|^2$$

Wiener filtering

$$\tilde{S}(\omega) = \underbrace{\frac{|X(\omega)|^2 - |\tilde{N}(\omega)|^2}{|X(\omega)|^2}}_{H(\omega)} X(\omega) = H(\omega) \cdot X(\omega)$$

$H(\omega)$ 를 이용하여 Linear-phase FIR filter를 구현 가능.

EPD (end point detection)

noise-only 구간이 0 ~ 0.5 \Rightarrow not always true

energy-based EPD를 하여 noise-only 구간을 구함.

이 구간에서 $|\tilde{N}(\omega)|^2$ 추정하여 차감.