Digital Signal Processing in the Frequency Domain

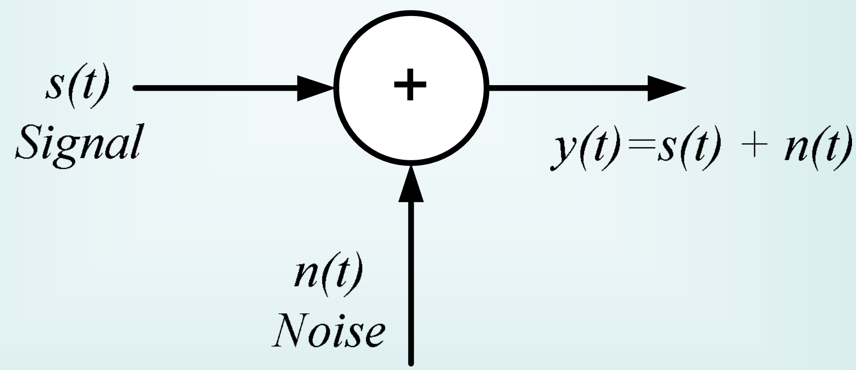
Dr Gilbert Soulodre, FAES

June 2018

camden labs

Noise Reduction

- A very common problem is having a desired audio signal that has been corrupted by noise
- We only have access to $y(t)$



- In general, we want to process $y(t)$ to reduce $n(t)$ and keep $s(t)$

Noise Examples

- Room noise, car noise
- Air conditioning (fan) noise
- Tape hiss, quantization noise
- Electrical noises (hums and buzzes)
- Tones
- Environmental noises (crowds, traffic)
- Clicks and pops (vinyl records)

Noise Reduction is Not Possible

$$y(t) = s(t) + n(t) \qquad s(t) = y(t) - n(t)$$

$$if \quad y(t) = 10, \ \text{what is } s(t)?$$

- If we look at the math, it is not possible to remove the noise from the signal.
- We do not know the signal
- We do not know the noise
- We only have a mixture of the signal and the noise

What if we us a truco?

- Can we find some trick that will make noise reduction possible?
- Yes!! if we take advantage of the properties of the ear.

Our Perceptual *Truco*

- The human auditory system is relatively insensitive to phase (phase differences or phase errors) over short periods of time (e.g. a few tens of milliseconds)
- From day-to-day experience we know that the ear is not overly sensitive to differences between the impulse responses within a given room.

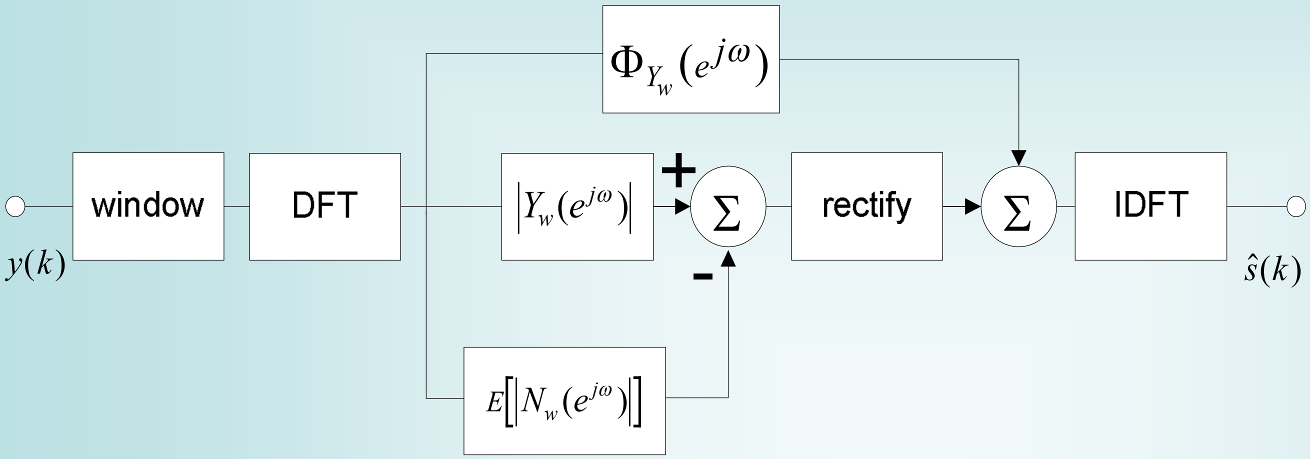
Our Perceptual *Truco*

- Therefore, provided that the processing is done over short blocks of time, the phase of the input signal can be used as a suitable estimate for the phase of the dry component of the signal
- *NOTE: The ear is NOT insensitive to phase over longer periods of time.*
- *Processing over long blocks (or the entire IR) will result in audible time-domain artifacts.*
- *Not the same as a phase error between two audio channels.*

Another truco….

- We assume that the magnitude of the noise does not change very quickly.
- This means we can use the average magnitude (spectrum) of the noise as our estimate of the noise in the whole audio signal.

Overview of Spectral Subtraction



$$Y_\omega = a_\omega + jb_\omega$$

$$|Y_\omega| = \left[a_\omega{}^2 + b_\omega{}^2 \right]^{1/2}$$

$$\Phi_{Y_\omega} = \arctan\left(\frac{\mathrm{b}}{\mathrm{a}}\right)$$

- method originally due to Weiss *et al*. and Boll
- assumes that noise is additive and not correlated to the signal
- assumes that noise is the same during speech and non-speech intervals
- the method relies on the fact that humans are relatively insensitive phase error over short time intervals
- permits reduction of both components of the camera noise

⇒ • performance is highly dependent on the SNR of the input signal ⇐

Derivation of Spectral Subtraction

$$y_w(k)=s_w(k)+n_w(k)$$ ⇐ signal windowed into frames (additive noise model)

$$\left|Y_w(e^{j\omega})\right|\approx\left|S_w(e^{j\omega})\right|+\left|N_w(e^{j\omega})\right|$$ ⇐ Boll's approximation

$$\left|\hat{S}_w(e^{j\omega})\right|=\left|Y_w(e^{j\omega})\right|-\left|N_w(e^{j\omega})\right|$$ ⇐ estimate of spectral magnitude of the clean signal

$$\left|\hat{S}_w(e^{j\omega})\right|=\left|Y_w(e^{j\omega})\right|-E\left[\left|N_w(e^{j\omega})\right|\right]$$ ⇐ spectral magnitude of noise approximated by expected value

$$\left|\hat{S}_w(e^{j\omega})\right|=\left|\left|Y_w(e^{j\omega})\right|-E\left[\left|N_w(e^{j\omega})\right|\right]\right|$$ ⇐ rectification

$$\Phi_{\hat{S}_w}(e^{j\omega})\cong\Phi_{Y_w}(e^{j\omega})$$ ⇐ use phase of noisy signal

$$\hat{S}_w(e^{j\omega})=\left|\hat{S}_w(e^{j\omega})\right|e^{\Phi_{Y_w}(e^{j\omega})}$$ ⇐ combine estimates of phase magnitude

$$\hat{s}(k)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{S}_w(e^{j\omega})e^{j\omega k}d\omega.$$ ⇐ inverse Fourier transform

Derivation of Spectral Subtraction

$$y_w(k) = s_w(k) + n_w(k)$$

$\Leftarrow$ signal windowed into frames (additive noise model)

$$\left|Y_w(e^{j\omega})\right| \approx \left|S_w(e^{j\omega})\right| + \left|N_w(e^{j\omega})\right|$$

$\Leftarrow$ Boll's approximation

$$\left|\hat{S}_w(e^{j\omega})\right| = \left|Y_w(e^{j\omega})\right| - \left|N_w(e^{j\omega})\right|$$

$\Leftarrow$ estimate of spectral magnitude of the clean signal

$$\left|\hat{S}_w(e^{j\omega})\right| = \left|Y_w(e^{j\omega})\right| - E\left[\left|N_w(e^{j\omega})\right|\right]$$

$\Leftarrow$ spectral magnitude of noise approximated by expected value

Derivation of Spectral Subtraction

$$\left|\hat{S}_w(e^{j\omega})\right| = \left| \left|Y_w(e^{j\omega})\right| - E\left[\left|N_w(e^{j\omega})\right|\right] \right|$$ $\Leftarrow$ rectification

$$\Phi_{\hat{S}w}(e^{j\omega}) \cong \Phi_{Yw}(e^{j\omega})$$ $\Leftarrow$ use phase of noisy signal

$$\hat{S}_w(e^{j\omega}) = \left|\hat{S}_w(e^{j\omega})\right| e^{\Phi_{Yw}(e^{j\omega})}$$ $\Leftarrow$ combine estimates of phase magnitude

$$\hat{s}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{S}_w(e^{j\omega}) e^{j\omega k} d\omega.$$ $\Leftarrow$ inverse Fourier transform

Generalized Spectral Subtraction

- various spectral subtraction algorithms vary their estimate of the spectral magnitude of the signal
- the spectral subtraction algorithm can be generalized as follows

$$\left|\hat{S}_W(e^{j\omega})\right|^{\gamma} = \left|Y_W(e^{j\omega})\right|^{\alpha} - \beta \cdot E\left[\left|N_W(e^{j\omega})\right|^{\alpha}\right]$$

where β is the overestimate parameter

$\alpha = 1,\ \gamma = 1$ $\Leftarrow$ Boll's method (magnitude subtraction)
$\alpha = 2,\ \gamma = 2$ $\Leftarrow$ power subtraction method
$\alpha = 2,\ \gamma = 1$ $\Leftarrow$ Wiener filter method

Zero Phase Filtering

- Derive on whiteboard

$$\tilde{S}_\omega = Y_\omega \bullet G_\omega$$

$$\tilde{S}_\omega = \left(a + jb\right)(c + jd\,)$$

$$\text{if G is real, } d = 0$$

$$\Phi_{Y_\omega} = \arctan\left(\frac{\text{bc}}{\text{ac}}\right) = \arctan\left(\frac{\text{b}}{\text{a}}\right)$$

- No change to the phase. Output phase is the same as the input phase.

Zero Phase Filtering

- Derive on whiteboard

$$\tilde{S}_\omega = Y_\omega \bullet G_\omega$$

$$\tilde{S}_\omega = (a + jb)(c + jd)$$

$$\text{if G is real, } d = 0$$

$$\Phi_{Y_\omega} = \arctan\left(\frac{\mathrm{bc}}{\mathrm{ac}}\right) = \arctan\left(\frac{\mathrm{b}}{\mathrm{a}}\right)$$

- No change to the phase. Output phase is the same as the input phase.

$$\tilde{S}_\omega = Y_\omega \cdot G_\omega$$

$$|\tilde{S}_\omega| = |Y_\omega \cdot G_\omega| = |Y_\omega| \cdot G_\omega \quad \text{---} \; 1$$

$$|\tilde{S}_\omega| = |Y_\omega| - |N_\omega| \quad \text{---} \; 2$$

use 1 & 2

$$|Y_\omega| \cdot G_\omega = |Y_\omega| - |N_\omega|$$

$$\therefore G_\omega = \frac{|Y_\omega| - |N_\omega|}{|Y_\omega|} = 1 - \frac{|N_\omega|}{|Y_\omega|}$$

Math Truco!!

Spectral Subtraction as a Zero-phase Filter

- the spectral subtraction process can be viewed as a zero-phase filter

$$\hat{S}_w(e^{j\omega}) = Y_w(e^{j\omega}) \cdot H(e^{j\omega})$$

- for example, setting $\alpha = 2$

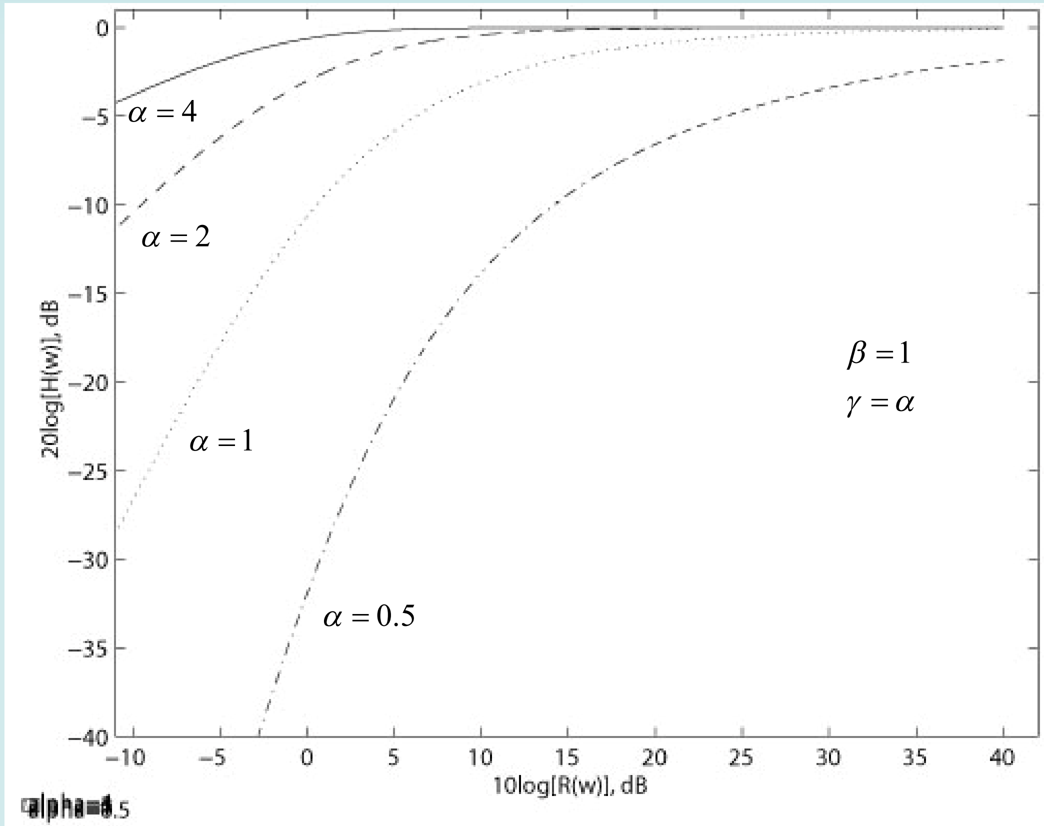
$$H(e^{j\omega}) = \frac{\hat{S}_w(e^{j\omega})}{Y_w(e^{j\omega})} = \left(\frac{\left|Y_w(e^{j\omega})\right|^2 - \beta \cdot E\left[\left|N_w(e^{j\omega})\right|^2\right]}{\left|Y_w(e^{j\omega})\right|^2} \right)^{1/2}$$

$$H(e^{j\omega}) = \left(\frac{X^2(e^{j\omega}) - \beta}{X^2(e^{j\omega})} \right)^{1/2}$$

where

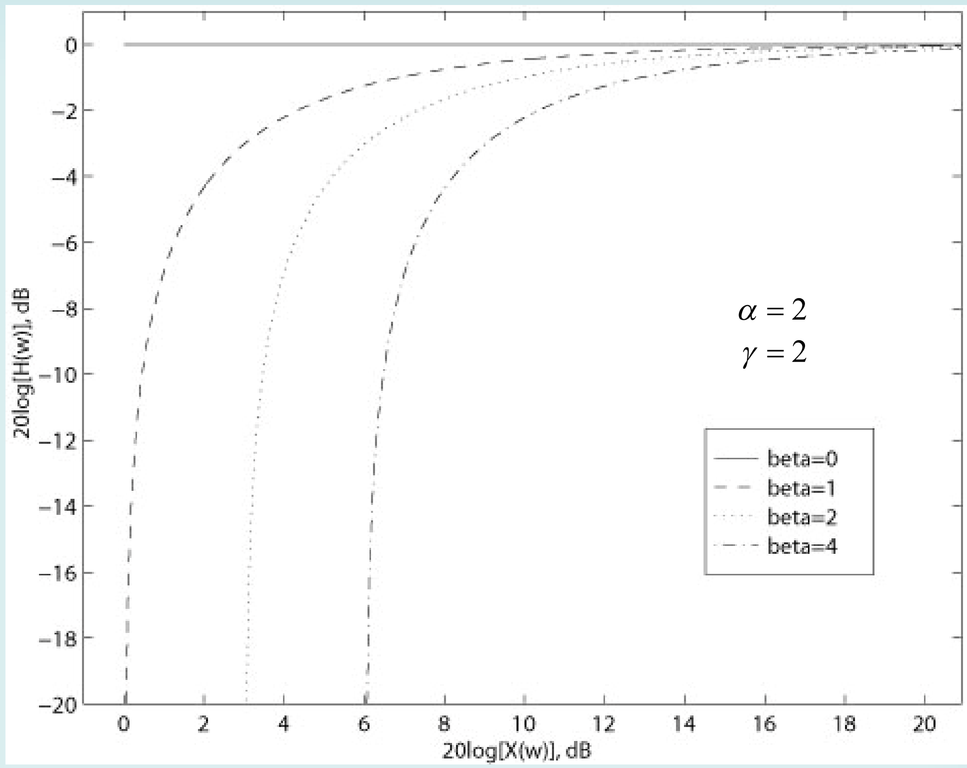
$$X^2(e^{j\omega}) = \frac{\left|Y_w(e^{j\omega})\right|^2}{E\left[\left|N_w(e^{j\omega})\right|^2\right]}$$

Spectral Subtraction Suppression Curves



- amount of attenuation depends on SNR of input signal
- shape of suppression curves determined by α, β, and γ

Spectral Subtraction Suppression Curves



- amount of attenuation depends on SNR of input signal
- shape of suppression curves determined by α, β, and γ

Limitations of Spectral Subtraction

- musical noise (most severe limitation)
- incomplete cancellation of noise (modulation of noise floor)
- timbral effects and loss of frequency components of the signal
- missing sounds - loss of low level signal components
- phase distortions
- temporal smearing

Steps must be taken to limit these artifacts.

Thank you!

camden labs