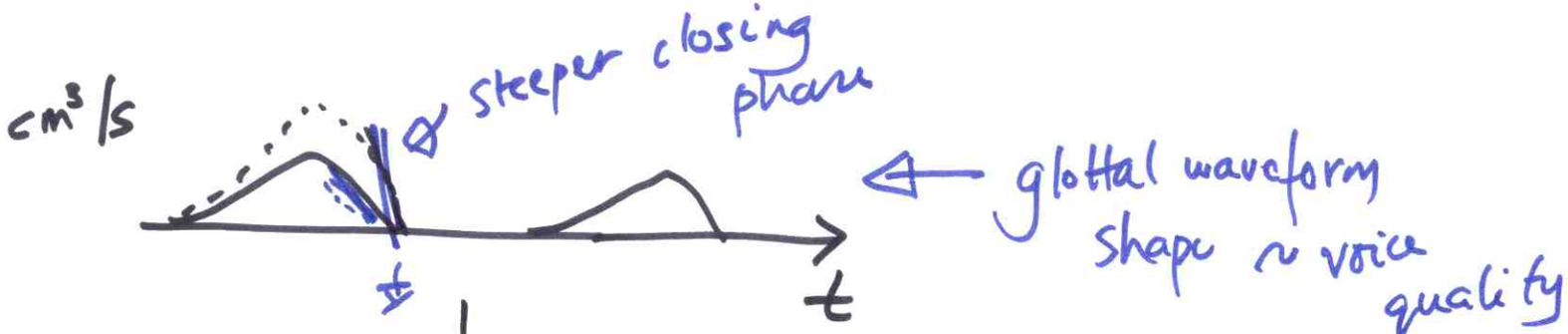
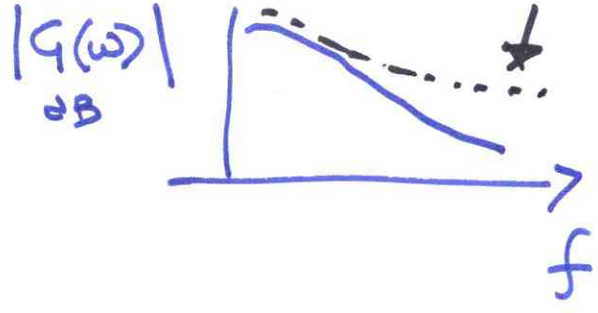
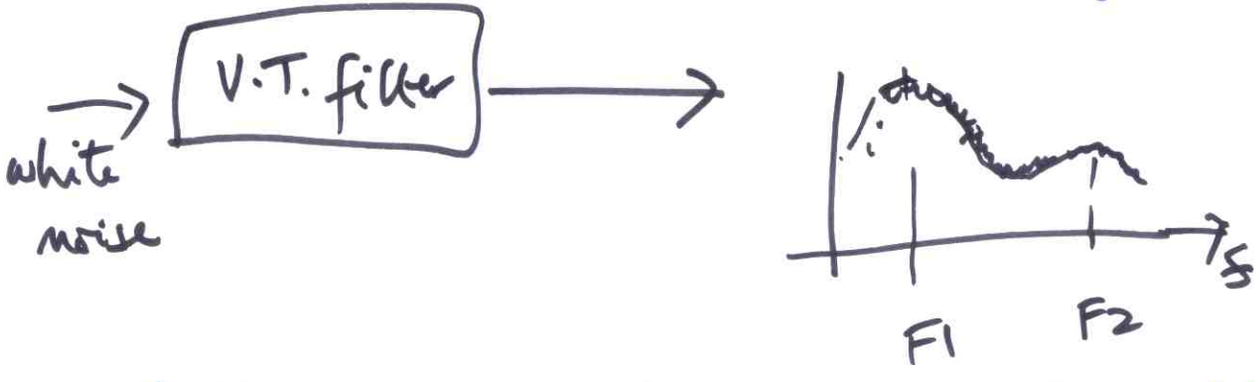






- Modal
- Breathy
- Hoarse
- Falsetto

production — Acoustics — Perception

· signal, spectrum

/m/, /n/

plosives

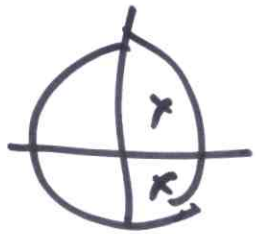
closure - release - transition into vowel

↑
burst of noise

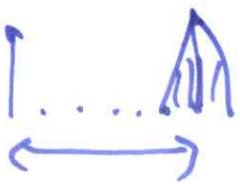
$y(n)$ $x[n]$ ← $H(z)$

↓

$|H(e^{j\omega})|$

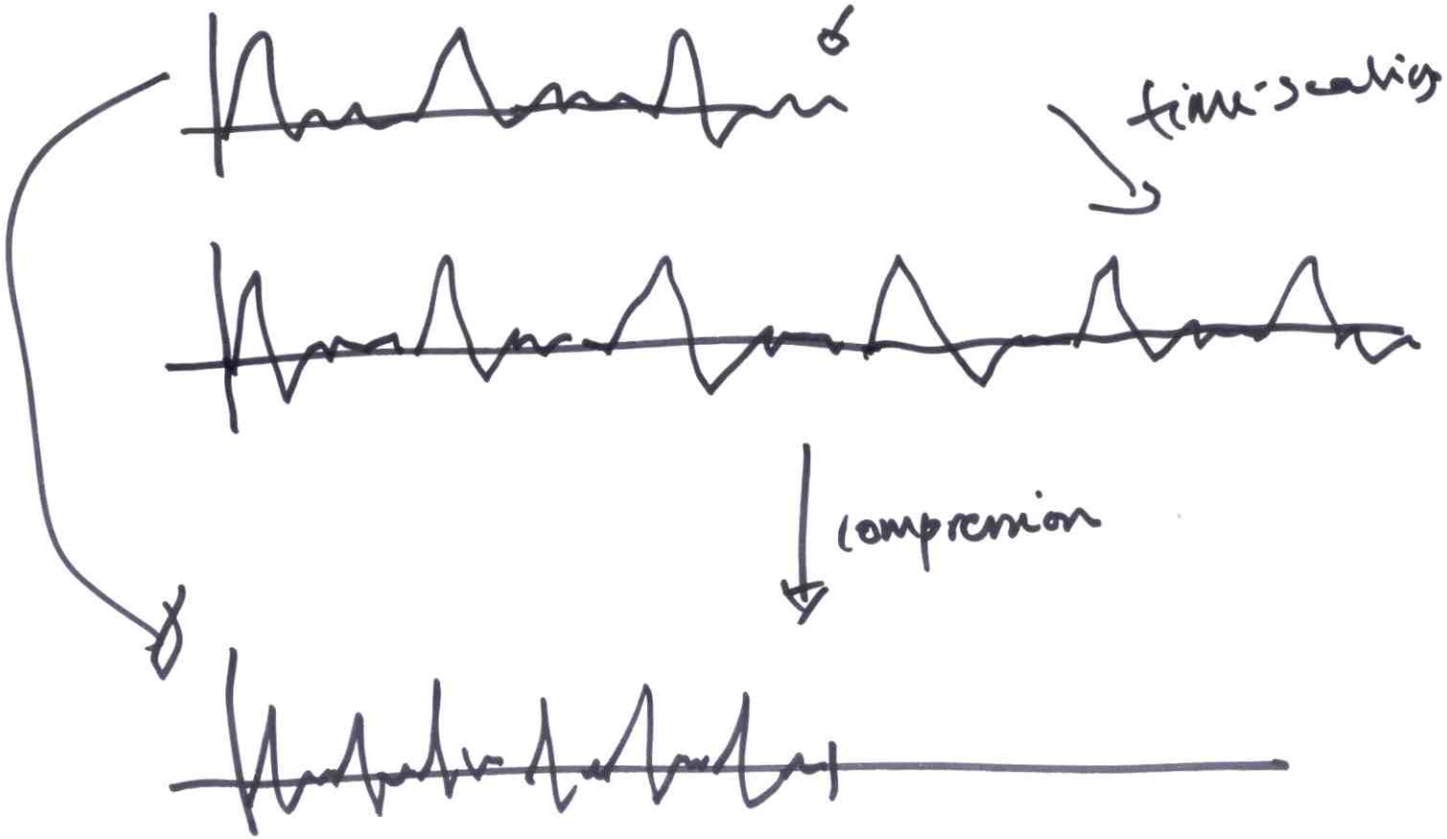


$x[n]$



$F_0 = 140$ Hz → $T_0 = \frac{1}{140}$ s.

Pitch-shifting



Time-domain methods of speech processing

1. Energy ← Short-time energy

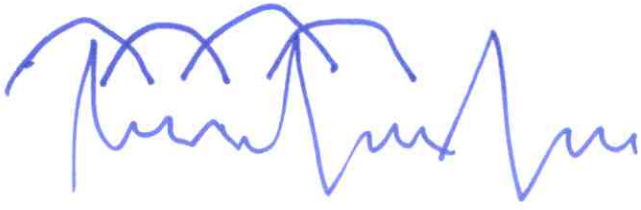
Any short-time analyses :

$$Q[n] = \sum_{m=-\infty}^{\infty} T\left(s[m]\, w[n-m]\right)$$

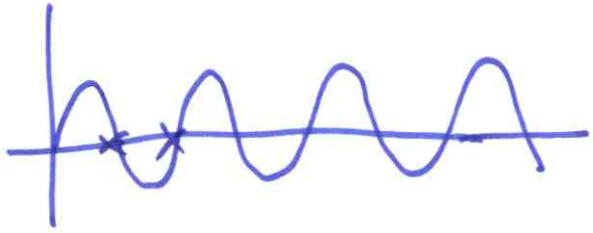
↑ fixed window

STE :

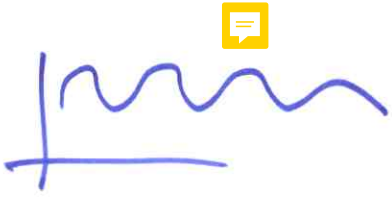
$$E[n] = \sum_{m=-\infty}^{\infty} s^2[m]\, w^2[n-m]$$

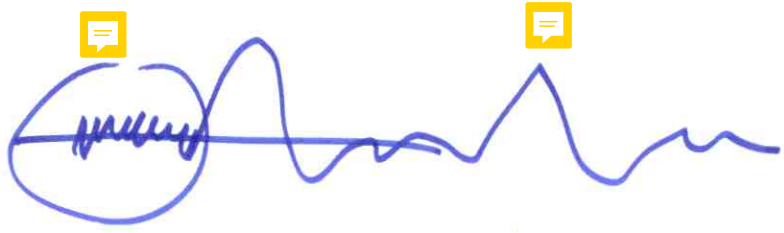


S-T ZCR



VAD



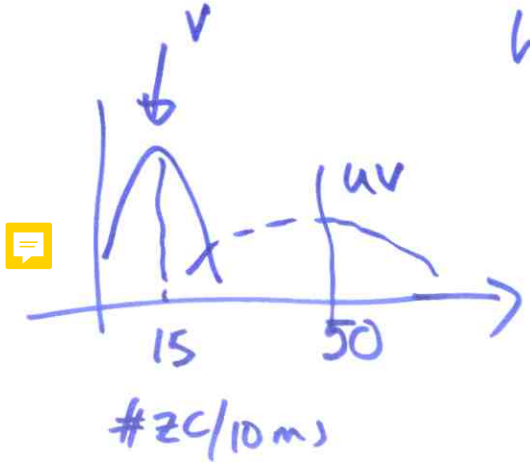


STE, ST ZCR

hi ZCR

low-ZCR

VAD



Speech or Silence