Pitch Detection

Nomenclature

- *pitch perception* vs. *pitch detection*
- *fundamental frequency* f_0 vs. *pitch*
- *buzz* vs. *hiss*
- pulses vs. noises

Difficulties in Pitch Detection

- large dynamic range of pitch: 60 Hz - 800 Hz
- pitch period fluctuate drastically and instantaneously
- vocal tract variation (e.g. /a m/)
- voiced-unvoiced transition
- telephone speech (convolutional noise)
- background noise

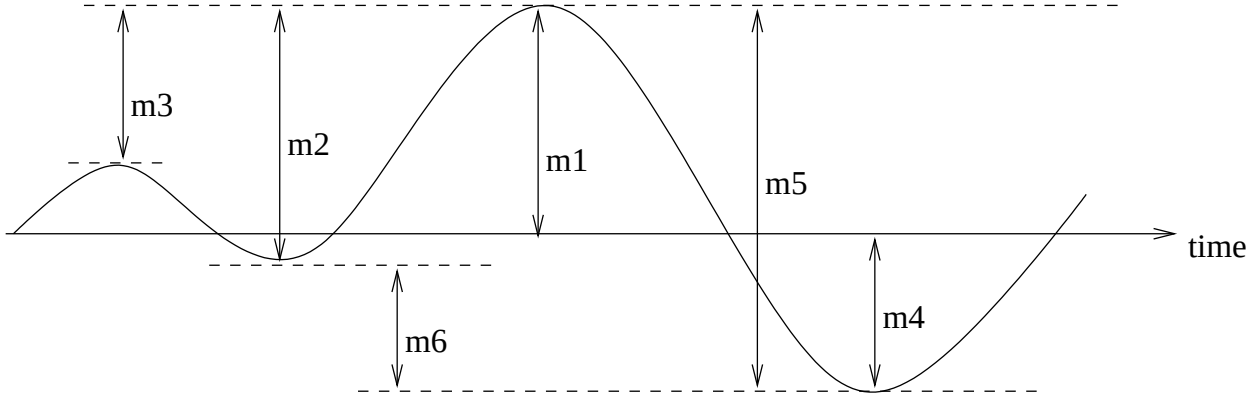
Ideas and Techniques

- **low-pass filtering**: it is easier to pick the peaks with smoothed signal
- **spectral flattening and correlation**: equal amplitude sum of harmonics results in pulse train
- **inverse filtering**: inverse of the vocal tract filter
- **comb filtering**: subtract delayed signal and look for minimum
- **cepstral processing**: high-time part is related to pitch period
- **high-resolution spectral analysis**

$$f_0 = \quad \text{the space between spectral lines}$$

A Multiple Information Source Method

- A low-pass filtering
- processor that extracts quantities from a speech window



- apply pitch period estimation from respective quantities
- combine the results
- post-processing: median smoothing