Pitch Detection

Introduction

Applications of Pitch Detectors

- Determine the excitation source of speech production.
- Pitch contour can be used for speaker recognition.
- Speech instruction for hearing impaired.
- Required by all speech analysis-synthesis (vocoder) systems.

Problems in Pitch Detection (Difficulties)

- Glottal wave is not truly periodic.
- Interaction between glottal excitation and vocal tract.
- Various ways to define the beginning and ending of each pitch period for voiced segment.
- Distinguishing between unvoiced and low-level voiced speech.
- Attenuation in telephone systems.

Types of Pitch Detectors

- Time-Domain Methods
- Frequency Domain Methods
- Hybrid Methods

Content

- Introduction
- Time-Domain Methods
  - Modified Autocorrelation Method Using Clipping
  - Parallel Processing Method
  - Data Reduction Method (陳建中)
  - Average Magnitude Difference Function
- Frequency-Domain Methods
- Hybrid Methods
  - Simplified Inverse Filtering Tracking (SIFT)
### Time-Domain Methods

- **Basic Assumption:**
  - Voiced speech is quasiperiodic
  - By minimizing the effects of the formant structure, simple time-domain measurements provide good estimates of the period.

- **Measurements**
  - Peak and valley measurements
  - Zero crossing measurement
  - Short-time energy and/or average magnitude measurements
  - Average magnitude difference measurement

### Frequency-Domain Methods

- **Fact:**
  - A periodic signal whose spectrum will consist of a series of impulses at the fundamental frequency and its harmonics.

- **Measurements**
  - Frequency Spectrum
  - Cepstrum

### Hybrid Methods

- **Using frequency-domain technique to provide a spectrally flattened time waveform, and then use autocorrelation measurements to estimate the pitch period.**
Pitch Detection

Modified Autocorrelation Method Using Clipping

To perform pitch detection, it is beneficial to remove the effect of vocal tract.

After removing, the spectrum is then flattened.

Spectrum Flattening

What method can be used to flatten the spectrum of the speech signal?

LPC is one of the candidates.

Whatever reasonable methods, in fact, can also be used although they are not directly correlated with vocal tract model.

Interpretation of Speech Waveform

Damping is mainly due to the vocal tract effect.

Therefore, we are trying to remove this effect.

Center Clipping

Evaluate the autocorrelation coefficients on the clipped signal.

Example
The Effect of Clipping Levels

AUTOC Pitch Detector

The Effect of Lowpass Filter

3-Level Center Clipping Function

It makes the computation of autocorrelation much easier.

Pitch Contours

Pitch Detection

Parallel Processing Method
PPROC Pitch Detector

- Pitch Extractor 1
- Pitch Extractor 2
- Pitch Extractor 3
- Pitch Extractor 4
- Pitch Extractor 5
- Pitch Extractor 6

Parallel Processing of Speech Signals

Silence Detector

LPF: 0-900 Hz

Peak and Valley Measurements

- Peak and Valley Measurements
- Peak and Valley Measurements

V/U Based on Agreement among 6 pitch detectors

Voiced Period = IP0

Unvoiced

Pitch Extractor 1
Pitch Extractor 2
Pitch Extractor 3
Pitch Extractor 4
Pitch Extractor 5
Pitch Extractor 6

m1 pulses
m2 pulses (peak-valley distance)
m3 pulses (peak differences; set negative value to zero)

Examples

Input (sinusoid) Input (weak fundamental and 2nd harmonic)

- The same fundamental period as the speech signal
- Different
Pitch Extractor

Variable Blanking Time (t) 
Variable Exponential Decay (e^-β)

\[ t \approx 0.4 P_{av} \]
\[ P_{av}(n) = \frac{P_{av}(n-1) + \beta P_{av} + \beta^2 P_{av} + \beta^3 P_{av}}{P_{av}} \]

Pitch Period Estimation

Unvoiced

Voiced

Pitch Detection

Simplified Inverse Filtering Tracking

Estimates of Period

\[ M_i \]
\[ \tau_0 \]

Pitch Period Estimation

Bias

1 2 5 7 10 12 15 20 30 50

Pitch Period Range (mae)

1.6~3.3

3.4~6.3

6.4~12.5

12.7~25.5

Coincidence Window Width in Hundreds of Microseconds.
Block Diagram of SIFT

Typical Signals from the SIFT Algorithm