Introduction to **Audio Content Analysis**

module 7.3.5: fundamental frequency detection — evaluation

alexander lerch

Georgia Tech | Center for Music Technology
College of Design

introduction

overview

Georgia Tech Center for Music Technology College of Design

corresponding textbook section

section 7.3.5

- **lecture content**
  - evaluation of pitch tracking systems
  - challenges in annotation
  - metrics

- **learning objectives**
  - successfully plan a systematic evaluation procedure for a pitch detection system

introduction
overview

corresponding textbook section

section 7.3.5

- **lecture content**
  - evaluation of pitch tracking systems
  - challenges in annotation
  - metrics

- **learning objectives**
  - successfully plan a systematic evaluation procedure for a pitch detection system

pitch evaluation
tasks

goal: compare predicted pitch and ground truth pitch

- differentiate various 'pitch tracking' tasks
  - pitch of **individual notes**
  - pitch of **monophonic melody**
  - pitch of **pre-dominant melody** in polyphonic mixture
  - pitches in multi-timbral **polyphonic mixture**

pitch evaluation
tasks

goal: compare predicted pitch and ground truth pitch

- differentiate various 'pitch tracking' tasks
  - pitch of **individual notes**
  - pitch of **monophonic melody**
  - pitch of **pre-dominant melody** in polyphonic mixture
  - pitches in multi-timbral **polyphonic mixture**

pitch evaluation
annotation challenges

- pitch discretization
  - (MIDI/score) pitch of individual notes
  - F0
- time discretization
  - start and stop time of note
  - equidistant time stamps

how to annotate F0

pitch evaluation
annotation challenges

- pitch discretization
  - (MIDI/score) pitch of individual notes
  - F0
- time discretization
  - start and stop time of note
  - equidistant time stamps

how to annotate F0

pitch evaluation

metrics — score ground truth

- all metrics should be computed in the pitch domain, not the frequency domain
- metrics measure a **match** between ground truth and predicted pitch ($\Rightarrow$ tolerance)
- **Raw Pitch Accuracy**:

$$RPA = \frac{\sum_{\forall n} TP_n}{\mathcal{N}}$$

$$TP_n = \begin{cases} 0, & \text{if } |\mathfrak{p}_{\mathrm{GT}}(n) - \hat{\mathfrak{p}}(n)| \geq 0.5 \\ 1, & \text{otherwise} \end{cases}$$

- **Raw Chroma Accuracy**:

$$RCA = \frac{\sum_{\forall n} TP_{\mathrm{chroma},n}}{\mathcal{N}}$$

$$TP_{\mathrm{chroma},n} = \begin{cases} 0, & \text{if } \mod\left(|\mathfrak{p}_{\mathrm{GT}}(n) - \hat{\mathfrak{p}}(n)|, 12\right) \geq 0.5 \\ 1, & \text{otherwise} \end{cases}$$

pitch evaluation

metrics — score ground truth

- all metrics should be computed in the pitch domain, not the frequency domain
- metrics measure a **match** between ground truth and predicted pitch ($\Rightarrow$ tolerance)
- **Raw Pitch Accuracy**:

$$RPA = \frac{\sum_{\forall n} TP_n}{\mathcal{N}}$$

$$TP_n = \begin{cases} 0, & \text{if } |\mathfrak{p}_{\mathrm{GT}}(n) - \hat{\mathfrak{p}}(n)| \geq 0.5 \\ 1, & \text{otherwise} \end{cases}$$

- **Raw Chroma Accuracy**:

$$RCA = \frac{\sum_{\forall n} TP_{\mathrm{chroma},n}}{\mathcal{N}}$$

$$TP_{\mathrm{chroma},n} = \begin{cases} 0, & \text{if } \mod\left(|\mathfrak{p}_{\mathrm{GT}}(n) - \hat{\mathfrak{p}}(n)|, 12\right) \geq 0.5 \\ 1, & \text{otherwise} \end{cases}$$

pitch evaluation

metrics — f0 ground truth

- all metrics should be computed in the pitch domain, not the frequency domain
- metrics measure **deviation** between ground truth and predicted pitch
- MSE, MAE, standard deviation from the ground truth

pitch evaluation

metrics — f0 ground truth

- all metrics should be computed in the pitch domain, not the frequency domain
- metrics measure **deviation** between ground truth and predicted pitch

- MSE, MAE, standard deviation from the ground truth

pitch evaluation
result aggregation

- aggregate per datapoint (frame/note)
- aggregate per file

summary
lecture content

- **potential data problems**
  - pitch and time quantization
  - reliability of ground truth
  - time resolution mismatch of ground truth and system
- **metrics**
  - score pitch match (chroma match)
  - measures of deviation
- **factor impacting metrics**
  - voicing detection