Beat Tracking based on Multiple-agent Architecture
— A Real-time Beat Tracking System for Audio Signals —

School of Science and Engineering, Waseda University
Masataka Goto
Yoichi Muraoka

1996/12/11 ICMAS 96

1. Introduction

1. Application of Multiple-agent Architecture
   - Application to beat tracking for musical audio signals
     Understand real-world audio signals
     Need to handle various ambiguous situations
   - Advantages
     Interpret input signals in various ways
     Multiple agents can examine multiple hypotheses
     according to different strategies
   - Main contribution
     Multiple-agent architecture is actually useful
     for a practical real-world application

2. What is Beat Tracking?
   - Track quarter notes
     just as people keep time to music by foot-tapping

3. Why is Acoustic Beat-tracking Important?
   - Computational model of human music perception
     Basic unit of the temporal structure of music
   - Useful in various applications
     Video/Audio editing systems
     Stage lighting control
     Music-synchronized CG animation

4. Beat Tracking Problem
   - Organize music into almost regularly spaced beats
     Musical audio signals
     (sounds of various kinds of instruments)
     Beat times
     (quarter-note level)
     Beat type
     (half-note level)
   - Assumptions
     Time signature: \( \frac{4}{4} \)
     Tempo: 61-120 M.M. almost constant

5. Issues in Tracking Musical Beats
   - Peak-finding with a threshold is not sufficient
     Many energy peaks are not directly related to beats
   - Multiple interpretations of beats are possible
     No specific sound directly indicates the beat position
   - Ambiguous situations
     Several events may correspond to a beat
     Different inter-beat intervals seem plausible
   - Context-dependent decisions using musical knowledge
     Whether a beat is strong or weak
     Which is the best interpretation

6. Our Solutions
   - Detect multiple tracking cues
     Onset times in several frequency ranges
     Chord change possibility
   - Maintain multiple hypotheses
     Each corresponds to a hypothetical interpretation

7. 2. Multiple-agent Architecture
   - Multiple Agents Examine Multiple Hypotheses
     Agents track beats according to different strategies
     Each maintains a beat-position hypothesis:
     Next beat time
     Beat type
     Inter-beat interval
     Predicted next-beat time

8. Definition of "Agent"
   1. The agent interacts with other agents
      to perform a given task
   2. The agent evaluates its own behavior
      on the basis of the input
   3. The agent adapts to the input
      by adjusting its own behavior
3. System Description (for music without drum-sounds)

- **Model**
  - Frequency spectrum
  - Onset components
  - Compact disc
  - Musical audio signals
  - Beat information
  - Frequency spectrum
  - Onset components
  - Compact disc
  - Musical audio signals
  - Beat information
  - Frequency spectrum
  - Onset components
  - Compact disc
  - Musical audio signals
  - Beat information

- **Make a Hypothesis**
  - Onset-time vector (Result of Frequency Analysis)
  - Beat time
  - Beat type
  - Agent-generated hypothesis
  - Beat time
  - Beat type

- **Interaction**
  - Each agent interacts with another agent to perform the beat-tracking task
  - 12 agents are grouped into 6 pairs
  - Paired agents predict next beat times cooperatively
  - Agent pair 1
  - Agent pair 2
  - Compensate for one of typical tracking errors

- **Interaction**
  - Interaction through prediction field
    - Expectancy curve:
    - when the next beat is expected to occur
    - Height of each local peak:
      - the next beat-position possibility
    - Paired agents inhibit each other's field

- **Evaluation**
  - 2. Each agent evaluates its own hypothesis according to the input acoustic signals
    - Reliability of the hypothesis
      - Quantitative result of the self-evaluation
    - Evaluate using musical knowledge
      - Manager decides which is the best hypothesis according to the reliability

- **Evaluation**
  - Two kinds of musical knowledge
    - 1. Sounds are likely to occur on beats
      - Correct beat times tend to coincide with onset times
    - 2. Chords are more likely to change on beats than on other positions
      - between two successive correct beats

- **Evaluation**
  - Manager
    - 1. Classify all hypotheses into groups according to beat time and inter-beat interval
    - 2. Calculate overall reliability of each group
    - 3. Select the dominant group
    - 4. Repeat three times while narrowing the allowable margin of beat times
      - the reliable hypothesis in the most dominant group

- **Adaptation**
  - 3. Each agent adapts to the current input by adjusting its own strategy parameter
    - The reliability becomes high enough
      - Tune a parameter to narrow the range of possible inter-beat intervals
      - Examine only a neighborhood of the current appropriate inter-beat interval
      - Keep the good hypothesis that has the inter-beat interval appropriate to the input
4. Experimental Results

☐ Conditions  (Implemented on AP1000)
  • Audio signals without drum-sounds from CDs
  • 40 popular songs performed by 28 artists
  • Tempo: 62-116 M.M. roughly constant

☐ Results
  • Correctly tracked beats in 34 out of 40 songs
  • Mistakes in 6 songs
    - beat times were wrong (4)
    - very few onset times
    - temporarily fluctuated tempo
    - beat type was wrong (2)
    - irregularity of chord changes

5. Conclusion

☐ Summary
  • Multiple-agent architecture for beat tracking
  • Examine multiple hypotheses in parallel
  • Each agent is capable of
    - interaction compensate for typical error
    - self-evaluation decide which is the best hypothesis
    - adaptation keep the good hypothesis
  • Robust enough to handle real-world audio signals

☐ Future Work
  • Upgrade our beat-tracking system
    to make use of other higher-level musical structure
  • More sophisticated interaction among agents
  • Application of the multiple-agent architecture
    to other perceptual problems