Offline Music-to-Score Following

Advisor: Prof. Shyh-Kang Jeng
Hermes Lyu

Dept. of Electrical Engineering, National Taiwan University

Motivation

- Keeping up with the music and the score simultaneously is difficult for amateur music lovers.
- Music to Score Alignment is the solution.
- A well aligned music-score pair can also be used as a music editing interface.

Related works

- 1984-1997: String matching techniques. Mainly use pitch
- 1997- : Statistical models(HMM model, Dynamic Time Warping, etc.)
- MIREX 2006, 2008, 2010 & 2011 (online only)

Multi-instrument Problem

- Most works focused on single or at most 3 instruments.
- Or they may require high computational cost.
- Melody alignment not suitable.

Experiment Setup

- Audio input:Record of Orchestral Music, Multi-instrumental
- Score input: MIDI
- Offline (i.e. not real-time)
- Accuracy: music appreciation purpose

System Flowchart

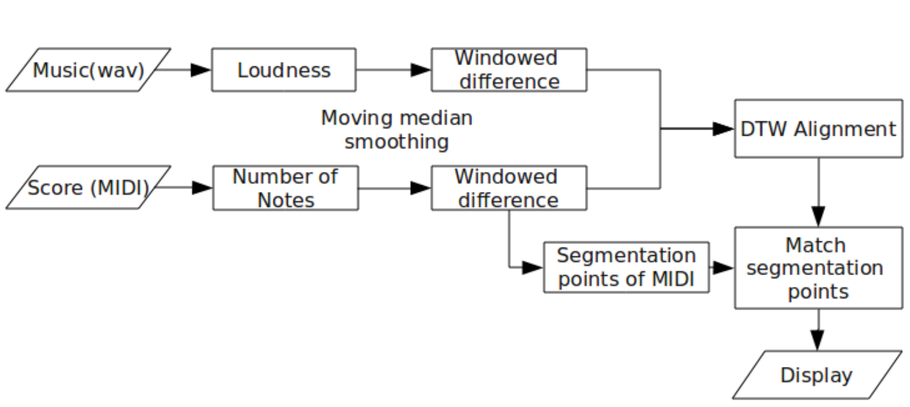
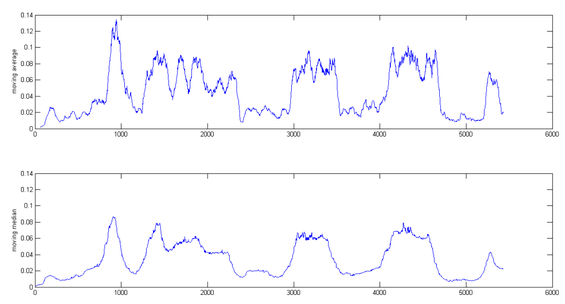


Figure: Flow Chart

Feature Extraction

- Music **loudness**
- **Number of notes** in the score
- Smooth with **moving median**



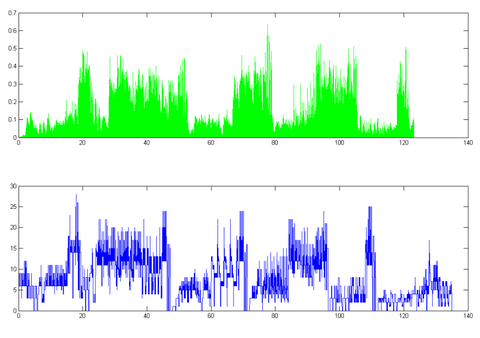


Figure: Green: wave energy, Blue: Number of notes

Figure: After moving median.

Segmentation

- Music tend to have segments.
- Significant loudness drop at segment boundary.(i.e. Quiet part after climax.)
- Tempo are roughly constant within each segment.
- By aligning the segmentation points, we can:
  - ignore the local error between two inputs.
  - get a acceptable accuracy for our purpose.

Windowed Difference

- Loudness dropping period have uncertain length.
- The dropping feature is hard to capture with simple difference.
- Windowed difference (window size = 2 sec)
  - Take two consecutive window.
  - Calculate the RMS average in each window.
  - Subtract the next window average with the previous window average.
- Negative peaks that are lower than 3 times of the average negative peaks are taken as segmentation points (MIDI only).

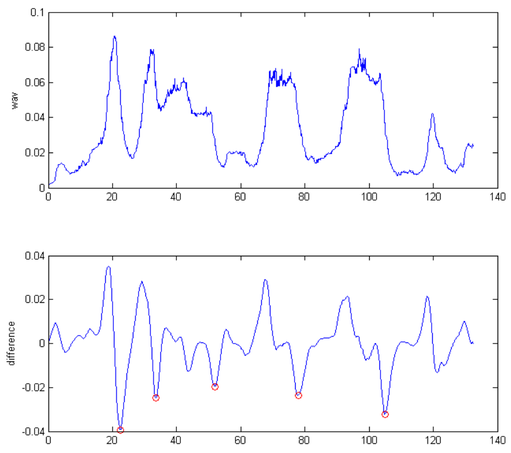


Figure: Features after windowed difference

Alignment - Dynamic Time Warping

- Find a minimal-cost matching relationship between two time series.
- Distance measure: Euclidean distance of data points.
- Construct a cumulative cost map.
- Find the minimal-cost path on the map.
- Project the segmentation points from MIDI to wave with this relationship.

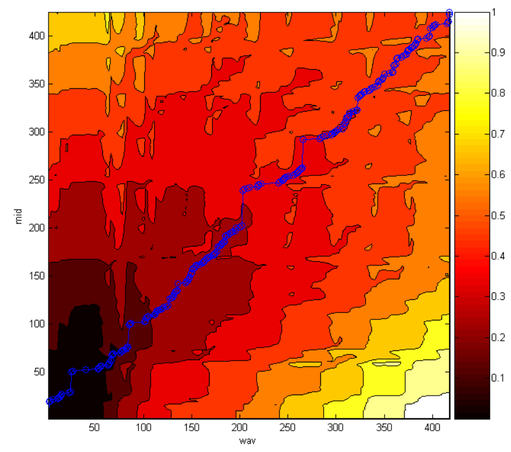


Figure: Cost map and path (blue line)

$$D(i,j) = |p(i) - q(j)| + min \begin{pmatrix} D(i-1,j), \\ D(i-1,j-1), \\ D(i,j-1) \end{pmatrix}.$$

Display

- Score are represented as a long continuous flow.
- A viewing window slides through the score. It's speed are decided by the segmentation points relationship.
- As long as the matching error is smaller than the window size, the viewer can still follow up.
- Piano-rolls are used. Easy to replace with traditional score.

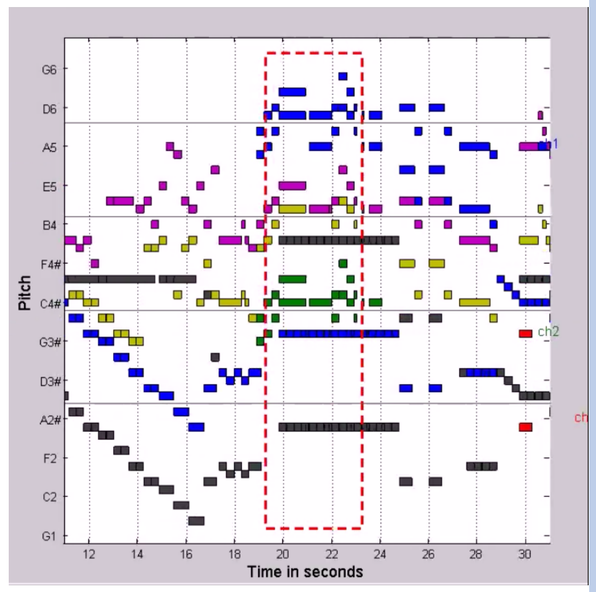


Figure: Color stands for different instrument.

Results and Future Works

- Error within ± 2 sec.
- Without alignment: ± 4 sec.
- Acceptable accuracy for appreciation purpose
- Improvements:
  - Different weighting for different instruments.
  - Switch to state-of-the-art single instrument methods for solo sections.
  - Use other score formats (e.g. MusicXML).
  - A more quantitative evaluation method.

National Taiwan University

Mail: b97901047@ntu.edu.tw