Semi-automated extraction of expressive performance information from acoustic recordings of piano music

Andrew Earis
Outline

- Parameters of expressive piano performance
- Scientific techniques: Fourier transform & wavelet transform
- CHARM: The semi-automated system
Measurable features of expressive performance

- **Timing (note onset and offset)**
- **Dynamics (variations in intensity of notes and chords)**
- **Pitch**
Timing parameterisation

- Note onset time. The time at which the note is played.
- Note offset time. The time at which the note is released.
- Attack duration. The time from the initial sound until its amplitude is at a maximum.
- Sustain duration. The time from maximum sound amplitude until the note is released.
- Decay time. The time from when the note is released until the sound ceases.
Dynamics parameterisation

- Power of note
- Maximum amplitude (amplitude of attack)
- Sustain amplitude (amplitude, or change of amplitude, of the sustain)
Pitch parameterisation

- Fundamental frequency of note
- Frequency of individual partials (and how they deviate from the precise harmonic frequencies)
- Vibrato
  - Frequency of vibrato
  - Amplitude of vibrato
- Glissando
Aims

- to extract the expressive performance information from acoustic recordings of piano music with reference to a digitised version of the musical score of the work being performed.
Digitised musical score

Acoustic recording

Note-on-note synchronisation algorithm

Event (note or chord) onset times dynamics

Individual note onset times offset times dynamics

Expressive performance data of individual performance
Methods

The location in time of individual notes and chords within an acoustic recording must be identified and related to a digitised score of the work being performed.

This is a classical time-frequency analysis problem.
A number of constraints on the time-frequency resolution required.

- **Required frequency resolution**
  - determined by the interval between adjacent semitones
  - proportional to the frequency of the lower semitone
  - Spacing varies from about 1.6Hz at the bottom of the piano keyboard to over 120Hz at the top
  - the fundamental frequencies ranging from 27.5Hz to just over 4kHz.

- **Required timing resolution**
  - The shortest note duration in rapid note passages is around 50ms.
  - The human ear is capable of perceiving differences of as little as 10ms.
Possible approaches

- Filterbank techniques
- Time-frequency analysis (using Fourier Transforms, wavelets and spectrographs)
- Modelling
Short Time Fourier analysis
Wavelet analysis

A wavelet spectrograph of the opening bar of Chopin Etude in E major. The colour scaling is in decibels.
Problems with ‘simple approach’

- Synchronization between score and acoustic recording relies on stepping through each event in the score in turn.
- If a particular note or chord in the score is not played, the whole system can break down.
- Other problems:
  - too much noise in the recording
  - a high degree of rubato
  - chords are played with too much asynchrony
- The longer the passage of music to be analysed, the more likely it is that such a system will break down.
- Time-consuming manual checking of the data would also required to ensure accuracy.
CHARM: Semi-automated system

- A semi-automated expression extractor, requiring user input for checking and verifying results as the algorithm proceeds i.e. note-on-note.
- Input to system: MIDI score, acoustic waveform
- Technology incorporated: filterbank/wavelet-based note onset detector
- Note/event parameter measurements: onset time, dynamics
Algorithm plan

STEP 1
Digitisation of musical score

STEP 2
Manual beat tapping to acoustic recording

STEP 3
Approximate performance score → Estimate of main beat onset times

STEP 4
Measurement of actual main beat onset times using algorithm

STEP 5
Manual correction of main beat onset times

STEP 6
Calculation of other event onset times and all event dynamics, individual note onset times / dynamics note onset timings and event/individual note dynamics
Advantages

- Quick to use, results can be verified as they go along – know they are correct!
- Very easy to use
- Wide range of scores can be used
- Quick and easy to develop and automate further
- Results are verified by the user as measurements are made, ensuring that all results are correct and can be used reliably for musicological analysis
- Further enhancements can be made (e.g. increase level of automation, sophistication of parameters)
STEP 1

- Digitisation of musical score
  - Score is scanned using a standard flatbed scanner
  - Converted to symbolic data with Sharpeye
  - Errors are corrected manually
  - Symbolic data is then converted to the Humdrum data format
  - This is outputted as ASCII tabular data
STEP 2

- Manual beat tapping to acoustic recording
  - Standard beat tapping procedures
  - backbeat.exe
  - average of 20 complete taps initially!
## STEP 3

- Averaged manual beat tapped data combined with the digitized musical score
  - Approximate performance score

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>abstime</td>
</tr>
<tr>
<td>2</td>
<td>duration</td>
</tr>
<tr>
<td>3</td>
<td>note</td>
</tr>
<tr>
<td>4</td>
<td>metric level</td>
</tr>
<tr>
<td>5</td>
<td>bar number</td>
</tr>
<tr>
<td>6</td>
<td>absolute beat</td>
</tr>
</tbody>
</table>

- abstime: Average absolute time in ms of beat tapped timings
- duration: Expected duration of note in ms based on score duration
- note: MIDI number of note pitch
- metric level: 1 = downbeat, 0 = beat, -1 = offbeat
- bar number: measured from beginning of work
Approximate performance score:
Mazurka in A minor op 7 no 2

<table>
<thead>
<tr>
<th>Time</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
<th>Value 4</th>
<th>Value 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1912</td>
<td>646</td>
<td>76</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2558</td>
<td>463</td>
<td>77</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3021</td>
<td>154</td>
<td>76</td>
<td>-1</td>
<td>1</td>
<td>1.75</td>
</tr>
<tr>
<td>3175</td>
<td>603</td>
<td>57</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3175</td>
<td>603</td>
<td>62</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3175</td>
<td>603</td>
<td>65</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3175</td>
<td>603</td>
<td>74</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3778</td>
<td>652</td>
<td>57</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3778</td>
<td>652</td>
<td>62</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3778</td>
<td>652</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3778</td>
<td>652</td>
<td>77</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4430</td>
<td>1111</td>
<td>77</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4914</td>
<td>627</td>
<td>60</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>4914</td>
<td>627</td>
<td>65</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
STEP 4

- Correction of ‘main beat’ onset times (from beat tapped times) using wavelet-based algorithm, programmed in MATLAB
Parameters

- Search window parameters
- Wavelet parameter - $k$
- Wavelet parameter - $LPF$
- Wavelet tuning
- Feedback analysis
Search window parameters

![Graphs showing mean deviation with different search window parameters.](image-url)
Wavelet parameters - $k$

**Chiu**

Mean deviation

- Deviation (ms)
- Mean deviation (outliers >100ms removed)

**Friedman**

Mean deviation

- Deviation (ms)
- Mean deviation (outliers >100ms removed)
Wavelet tuning
Feedback analysis

Mean deviation

Mean deviation (outliers >100ms removed)

\[
a = [-0.33 \text{ IOI : 0.33 IOI}]
\]

\[
b = [-0.50 \text{ IOI : 0.50 IOI}]
\]

\[
c = [-0.67 \text{ IOI : 0.67 IOI}]
\]

\[
d = [-0.33 \text{ IOI : 0.67 IOI}]
\]

\[
e = [-1.00 \text{ IOI : 1.00 IOI}]
\]
STEP 5

- Manual correction of main beat onset times
  - Sonic Visualiser
- Output – table of corrected main beat onset times
STEP 6

- Estimation of subbeat onset times from main beat onset times
- Correction of estimated subbeat onset times using algorithm
- Estimation of main beat and subbeat dynamics
- Estimation of individual note onset times
End product

- MATLAB code converted to standard .exe files
- Input data required
  - Excel control spreadsheet - for batch processing of data
    - Input/output file names
    - Analysis parameters
  - ‘Main beat’ beat tapped input data spreadsheet
  - .wav sound recording
End product

- Three different algorithms
  - mainbeatcorrection.exe
    - correction of ‘main beat’ tapped onset times
    - measurement of main beat dynamics (in dB)
  - subbeatmeasurement.exe
    - measurement of sub beat onset times
    - measurement of sub beat dynamics
  - individualnotemeasurement.exe
    - measurement of individual note onset timings and dynamics

- Will be available (or a variation of this) on CHARM website as downloadable resource
End product

- For example...
  - 3 recordings of Chopin’s Mazurka in A minor op 7 no 2
    - Friedman (1930)
    - Rubinstein (1939)
    - Indjic (1988)
  - Graphs of
    - inter-onset interval for each event in score
    - dynamics of each event (scaled in dBs)
Inter-onset interval

Friedman (1930)
Rubinstein (1939)
Indjic (1988)
Event dynamics

- Friedman (1930)
- Rubinstein (1939)
- Indjic (1988)
Conclusions and future work

- Further scientific development (wavelet parameters, shape, dynamics measurement)
- Extension to note offset times
- Testing and refinement of user interface