Source Codes for Polyphonic Rhythm Transcription and Note Value Recognition

Eita Nakamura

2017/3/18

Contents

1 Package Content

- Rhythm transcription based on metrical HMM (`RT_MetricalHMM`)

- Note value recognition based on Markov random field model (`NVR_MRF`)

  For the rhythm transcription algorithms, the input and output files are given as *Trx files*, a special file format to describe transcription data (see Section 2.2).

- MIDI to Trx converter (`MIDIToTrx`)

- Trx to quantised MIDI converter (`TrxToQuantisedMIDI`)

  The converter `MIDIToTrx` inputs a standard MIDI file (SMF) and outputs a Trx file, making it ready to be processed by the transcription algorithms.

- Evaluation tool (`NVEvaluationTool`)

  The evaluation tool takes two Trx files, typically a ground-truth and an estimation, and calculates the error rate and scale error.

2 Compiling and Usage

2.1 Compile

To compile, process
$./compile.sh
or process lines written in this shell script file with adaptations to the environment.

2.2 File Formats

The package deals with two file formats, *standard MIDI file (SMF)* format and the *Trx* format. A 'MIDI file' in this manual always means an SMF.

The Trx format describes both the performance data written in a MIDI file and the score data including the onset and offset score times of each note. The first lines of an Trx file look like the following:

```
//Version: Trx_v170203
//TPQN: 4
//Meter: Duple
..... (Some comment lines)
0 0.01 0.12 C4 19 80 0 0 2 1 0.82 1 0 0.45
1 0.02 0.15 C5 32 80 0 0 2 0 0.82 0 0 0.45
.....
```

The second line describes the 'tick per quarter note (TPQN)', the integer corresponding to a unit of quarter note, which is used to describe score times. For example, if TPQN is 4, a 16th note has a value of 1 and a dotted half note has a value of 12. The third line indicates the metre: If it is `duple` then the metre is 2/4 and if `triple` then 3/4.

For the main content, each line describes a musical note. From left to right, columns indicate:
ID (onset time) (offset time) (spelled pitch) (onset velocity)
(offset velocity) (channel) (onset score time) (offset score time)
(voice) (local tempo) (ID1) (ID2) (end time)
The offset time is the key-release time given as note-off event in MIDI files and the end time is an offset time extended according to sustain and sostenuto pedal information. For the output of this package the voice is always 0 and `ID1` and `ID2` are always −1. The local tempo is described as 'second per quarter note'.

2.3 Usage

To run the programs, enter the following:
$./run.sh ex
Here 'ex' is the name of the MIDI file to process (`ex.mid`) and this should be replaced when running other files. This produces, the Trx converted from the input MIDI file (`ex_raw_trx.txt`), the result of onset rhythm transcription by the metrical HMM (`ex_onset_trx.txt`), the result of onset and offset transcription by the Markov random

field model (`ex_MRF_trx.txt`) and a quantised MIDI file (`ex_qunatised.mid`). Please see lines in the shell script file `run.sh` to separately run each algorithm.

Finally, given a ground-truth data `ex_groundtruth_trx.txt` and an estimated result `ex_MRF_trx.txt`, the evaluation tool can be run as follows:

`$./Programs/NVEvaluationTool ex_groundtruth_trx.txt ex_MRF_trx.txt`