

Programming Codes for Polyphonic Rhythm Transcription

Eita Nakamura

2016/12/20

Contents

1	Package Content	1
2	Compiling and Usage	2
2.1	Compile	2
2.2	File Formats	2
2.3	Usage	2

1 Package Content

- Rhythm transcription based on merged-output HMM (`RT_MergedOutputHMM`)
- Rhythm transcription based on note HMM (`RT_NoteHMM`)
- Rhythm transcription based on metrical HMM (`RT_MetricalHMM`)

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

- MIDI to Transcr converter (`MIDIToTranscr`)

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

- Evaluation tool (`CompareTranscrMultiVoice`)

The evaluation tool takes two transcr files, typically a ground-truth and an estimation, and calculates the edit distance (rhythm correction cost).

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 *transcr* format. A ‘MIDI file’ in this manual always means an SMF.

The *transcr* format describes both the performance data written in a MIDI file and the score data including the onset score time and the note value of each note. The first lines and the last line of a *transcr* file look like the following:

```
// TPQN: 4
// log prob 0
0 0.731 0.991 A4 66 80 1 0 1
1 0.736 1.188 A3 50 80 0 0 2
.....
end 14.029 64
```

The first line describes the ‘tick per quarter note (TPQN)’, the integer corresponding to a unit of quarter note. It is used to describe note values (score written quantised duration of notes) and score times (the score position written in terms of note values). For example, if TPQN is 4, a 16th note has a note value of 1 and a dotted half note has a note value of 12.

Starting from the third line, each line describes a musical note. From left to right, columns indicate:

```
ID (onset time) (offset time) (spelled pitch) (onset velocity)
(offset velocity) (channel) (score time) (note value)
```

The last line describes an onset time and the corresponding score time for the last note.

2.3 Usage

To obtain a *transcr* file from a MIDI file run `MIDIToTranscr` as follows:

```
./MIDIToTranscr in.mid out_transcr.txt
```

The rhythm transcription algorithms can be run as follows:

```
./RT_MergedOutputHMM in_trascr.txt out_transcr.txt
```

```
./RT_NoteHMM in_trascr.txt out_transcr.txt
```

```
./RT_MetricalHMM in_trascr.txt out_transcr.txt
```

Given a ground-truth `true_transcr.txt` and an estimation `est_transcr.txt`, the evaluation tool can be run as follows:

```
./CompareTranscrMultiVoice true_transcr.txt est_transcr.txt
```

The output is like `est_transcr.txt 109 9.0001 0.0825697`

where from left to right, the estimation input, the number of notes, the rhythm correction cost and the rhythm correction rate are indicated.