Task and Models

In this article, we present two models tackling note-wise music alignment, i.e., alignment of a MIDI performance with a corresponding musicXML score by means of matched pairs of notes. This type of alignment can happen in two ways: offline (with access to the full recording of the performance) or online (following the performance as it happens, possibly in realtime).

The two models are:
- a two-step dynamic time warping (DTW)-based offline model
- a reinforcement learning (RL)-based online model

The models share the design approach of completely separating the handling of pitch and timing information.

Offline DTW model

First DTW step:
Pitch and pitch-set based features to create a coarse mapping from performance to score.

Using this mapping, each score onset is projected to an approximate performance time.

Second DTW step:
All notes are separated into pitch-wise sequences. The score notes of each pitch are mapped to performance time using the approximation of the first DTW path. Then, a DTW is computed between the score onsets (mapped to p. time) and the performance time, for each pitch and using the time delta between onsets as local metric. This DTW path minimizes the global distance between the sequences. For each note which is connected to multiple reference notes (horizontal or vertical steps in DTW matrix), we match the notes with minimal distance, marking the other insertions and deletions, respectively.

Offline Evaluation

The offline model is not tuned and hence is evaluated on the whole of four high-quality datasets of note-aligned solo piano music.

The proposed model outperforms the previous state of the art significantly on all datasets except for Vienna4x22, where the state of the art already reaches a perfect F-score for many performances (see Table 1).

Online RL model

We (partially) formalize the online note alignment task as reinforcement learning problem. Specifically, we define an agent as moving on the score. The local score context, i.e., 8 score onsets before and after the agents, as well as the current performance context, i.e., the last 8 performed onsets, are given to the agent as current state (S) of its environment. The agent can pick from 16 actions (A), i.e. move to one of the score onsets within the context window. Only the choice of the score onset that corresponds to the most recent performed note receives a positive reward, all other actions are not rewarded.

We learn the agent’s state-action value function Q(S,A) using a small attention-based neural network. We design the agent to be completely myopic, i.e., the value associated with each state-action tuple corresponds only to the immediate reward, not any delayed information. We further sample possible contexts from a dataset which turns the reinforcement learning problem into an offline one with supervised value function training.

Online Evaluation

The online model is evaluated on an unseen test set of five pieces in two ways:
- as a score follower predicting the current score onset (Table 4), measured in asynchrony
- as a note alignment model using the same F-score used in the offline case

Implementations:

https://github.com/sildater/parangonar
pip install parangonar

References: [1] [2] [3]