Report: Jingju style music generation using Feedforward VAE

Huicheng Zhang, u188840

Motivation & Objectives

Chinese traditional music is fascinating: the beauty of pentatonic scale, the rich potential of expressing emotion through only melody line (so different than western harmony tradition). I want to build a deep learning system that can automatically generate Chinese traditional music. Luckily there is a symbolic dataset of Jingju music created by Rafael Caro and it is good starting point for studying Chinese folk music.

Jingju Musical System:

Jingju is a crystallization of collective composing dating back to Qing Dynasty, which can be viewed as an typical example of Chinese traditional music system. Here is several key concepts of Jingju:

Banshi - Rhythmic patterns:

- Kuaiban(Allegro): with short duration and dense notes, 1/4 meter, expressing happiness and nervousness.
- Manban(Adagio): Less notes and longer duration, 4/4 meter.
- Yuanban: Moderate, 2/4 meter

Shengqiang - Melodic framework that jingju music evolves: "characteristic pitch progressions". For example: Xipi, Erhuang

Role types - Modes, pitch range. For example: Sheng, Dan, Jing, Chou

Jingju Dataset Introduction and preprocessing:

- Banshi: BPM information is normalized to 120 while meter information is preserved within the socres.

- Shengqiang and role type label annotation is preserved in the filename. However, since all notes are transposed to E major, the actual key information of different shengqiang and role type is missing.
- A fixed length slicing window was applied to each score to create every single training example and then randomization was applied. This operation leads to the absence of high level music structure information
- After carefully observation of the symbolic dataset, parameters are chosen as follows: `lowest_pitch = 48 # C3` `n_pitches = 24 # Two Octaves` `len = 8 # every training sample has 8 beat`

Neural Network:

- 4 layer feedforward VAE was built, from 64x24 dimension to 64 dimension latent vector. Training parameter can be seen in the colab notebook.

Result Analysis:

- Anhemitonic Pentatonic Scale feature is learnt. Notes that are not in E major are seldom presented in the output. Coloration notes are (4th degree and 7th degree) occasionally presented which is supposed to be.
- Different Rhythmic patterns are captured: *Kuaiban(Allegro) with short duration notes, usually in a 16th or 32th notestyle; Manban(Adagio) with long duration notes, usually in a quarter note style.*
- Note B (5th degree in E major) was more frequent, and is served as the ceiling of musical range. This result is supported by the ethnomusicology study of Jingju.
- A sense of counterpoint: By setting lower threshold more notes are revealed and the generated music contains multiple melodic lines which are related somehow. This indicates hidden accompaniment was learnt.
- Every eight beat is one basic unit of music. The neural network did not learn the high level structure of jingju music (the couplet form) due to the fixed length of training sample.

Future Improvements:

- Neural network structure: From feed forward neural networks to different network structures. For example, BLSTM, in order to learn a high level structure of music.

- Future Improvements:
  - Music generation with User interaction(e.g. generating music with specified rhythmic pattern, Shenggqiang).
  - Output the musicxml file to check if the system compose counterpoint.
  - Different training scheme: To generate music given the motif by user.