MUSIC-SCORE ALIGNMENT IN PERFORMED CLASSICAL MUSIC

By

MinglangTuo

1508978

Supervised By

Shengchen Li

Project Specification

2021/9/25

1. PROBLEM STATEMENT

1.1 Problem

During my postgraduate study, I'm interested in the optimization of reinforcement learning. When I browsed the relevant literature, I found the problem of scores following music is attractive. In addition, it's about the application of reinforcement learning in audio processing. We will study how to use appropriate algorithms to design agents, so that agents can process audio efficiently in different environments.

As for scores following music [1], it's essentially a process of music score tracking. When the music is playing, the machine will automatically identify the music to draw the required notes and score them. The principle of the problem is the continuous control problem of reinforcement learning. It is different from the discrete control problem, which are only a few directions to explore. The most typical continuous control problem is the manipulator problem[2]. To solve this kind of problem, we can usually use two different strategy methods, one is based on the determined strategy gradient, the other is based on the random strategy gradient. Music score tracking also is a continuous control problem, which adjusts its reading speed through the current listening and playing performance.

1.2 Literature review

For solving this problem, I have browsed the relevant literature. At the beginning, people use traditional solutions, like manual creation or optical music software creation [3]. Traditional methods use fractional symbolization and computer-readable representation, similar to music-xml and MIDI. People can represent manually or use optoelectronic recognition software. However, with the rapid development of machine learning, it is found that convolutional neural network can be used to solve the problems identified by computer[4]. Nevertheless, there are limitations that can only identify a single channel, cannot deal with complex music problems (such as orchestral music) and cannot set a score evaluation mechanism. In recent years, with the development of reinforcement learning, it is found that reinforcement learning (DQN,Sarsa) can be used to train models to solve this problem. However, the comparison between two deep reinforcement learning algorithms is rarely mentioned in the literature, so we choose Proximal Policy Optimization and the Synchronous Advantage Actor Critic[5]. Our hypothesis is that Synchronous Advantage Actor Critic algorithm has more accurate scores than Proximal Policy

Optimization algorithm in training the same track music. If the conjecture is correct and follow-up time is sufficient, we try to change the Synchronous Advantage Actor Critic algorithm to make it more accurate and efficient than the original algorithm.

2. RESEARCH METHOD AND MATHEMATICAL PRINCIPLES

2.1 Research Method

For data collection, we will use the secondary method which means using the existing data of the website including Nottingham database(https://ifdo.ca/~seymour/nottingham/nottingham.html) and the MSMD dataset(https://github.com/CPJKU/msmd) to train the models. Nottingham dataset contains single channel music performance, while MSMD dataset contains more complex music, such as symphony. Taking the data set as the input, the corresponding relationship between each pixel position and each recording position is established to generate the corresponding digital score. Then we will establish relevant baselines and use a quantitative analysis to compare the score tracking of the two algorithms

2.2 Mathematical Principles

We will use the mathematical model of Markov chain[6], which can represent the transition between States, which means the interaction between agents and environments in different states. It can be represented by the following picture(Figure1). Each state shows the notes related to the current agent listening to the performance.

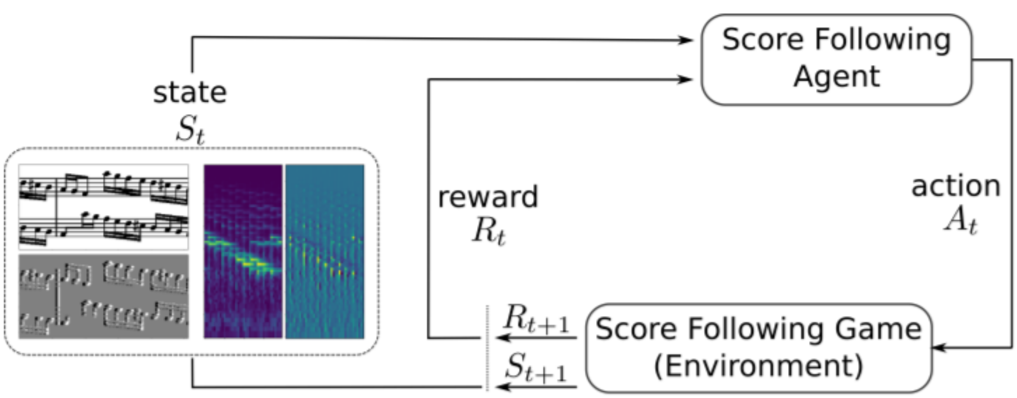

Figure 1

We will understand some basic signal conversion processes and use **fluidsynth**(Figure2). The software can be used to convert human voice features into forms that can be processed by computer.

Figure 2

After that, we must understand how to build the most basic neural network, because both policy based reinforcement learning and value based reinforcement learning need to be used in the follow-up. Especially for the later, we try to use the actor critic method, which will be built in combination with these two forms.

3. AIMS AND OBJECTIVES

3.1 Aims and Objectives

Music score tracking is the process of tracking music performance (audio) according to the known symbol representation (music score). The task of automatic score tracking system is to track the music performance related to the known symbol representation, that is, score. Music score tracking belongs to music retrieval, which itself has many subsequent application bases, such as automatic page turning, automatic accompaniment and visual synchronization of live music in concert. For the problems we want to study at this stage, we are divided into three steps:

1. Try to train the creation of multi-mode RL agent (environment configuration), and design score evaluation criteria, so that the agent can learn to listen to music at the same time in an end-to-end way, read the music score from the music score image and follow the audio in the music score, so as to obtain the maximum score of music score tracking as much as possible.

2. We intend to use different experiments to prove which reinforcement learning algorithm to train agents. We will use the most advanced algorithms, such as the dominant actor critic algorithm and Proximal Policy Optimization algorithm based on deep learning.

3. At the same time, we intend to analyze the advantages and disadvantages of the system, refer to relevant new literature, propose our own improvement method, and compare it with the latest advantage algorithm.

4. EXPECTED OUTCOMES AND PROJECT PLAN

4.1 Expected Outcomes

As for evaluation criteria, we found two different schemes from the literature. The first is the method described by dorfer [7], which models the following scores as a multimodal positioning task, divides the drawing segments into discrete buckets, and trains the neural network in a supervised manner to predict the most likely buckets under the current audio excerpt. The difference between the second is that it needs to preprocess the image. For the second criteria, we apply online dynamic time warping, namely optical music recognition (OMR), which converts the scanned music score into a computer-readable format, such as MIDI [8].

For the first one, it is relatively easy to judge. Directly take the image with the highest score in the real world and compare and judge it with the agent's image. The second is more complicated. We use music score to convert it into midi. In order to keep the extracted score consistent with the performance, we use fluidsync to synthesize MIDI data, 6 extract chromaticity features. According to the comprehensive score and performance audio, we use ODTW to align it, and judge whether it is good or bad by comparing these parameters.

In conclusion, no matter which method, we hope that the synchronous advantage actor critical algorithm can get higher scores than PPO method.

4.2 Project Plan

The following figures shows the all tasks about final year project. Now, the specification and design of project have been finished. While the programming work is still executing, which is expected to finish on 20th, Oct, 2021. And all other work will continue.

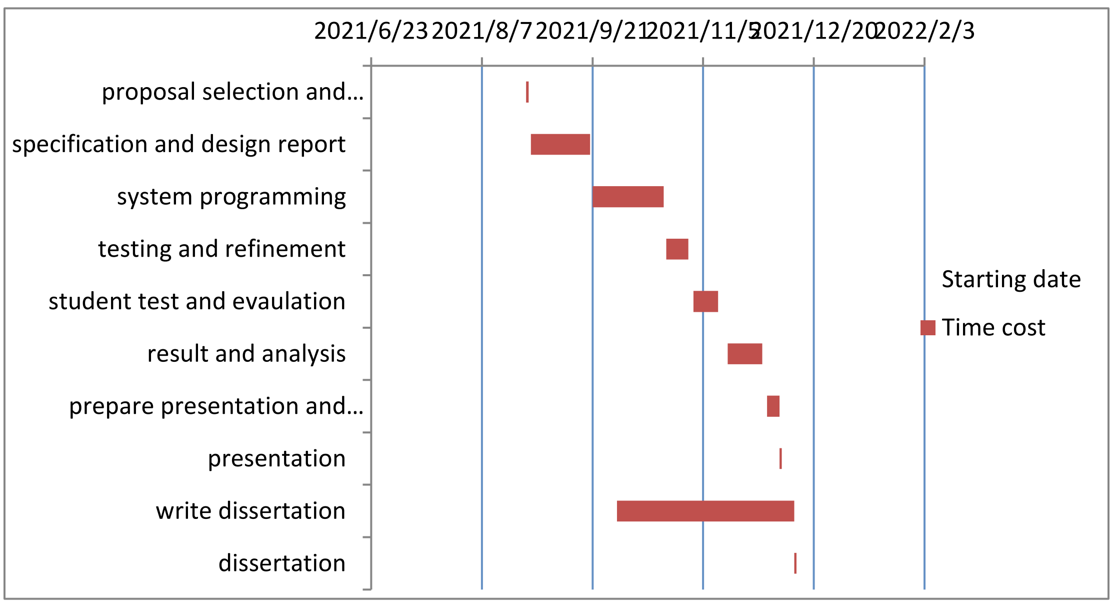

Figure 3

Task	Starting date	date of cost	Completing date
proposal selection and supervisor allocation	08/25/2021	2	08/26/2021
specification and design report	08/27/2021	23	09/20/2021
system programming	09/21/2021	30	10/20/2021
testing and refinement	10/21/2021	9	10/30/2021
student test and evaluation	11/1/2021	10	11/11/2021
result and analysis	11/15/2021	14	11/29/2021
prepare a presentation and write dissertation	12/01/2021	5	12/06/2021
Presentation	12/6/2021	1	12/7/2020
write a dissertation	10/1/2021	72	12/12/2021
Dissertation	12/12/2021	1	12/13/2021

Figure 4

REFRENCES CITED

[1] M. Dorfer, F. Henkel, and G. Widmer, “Learning to listen, read, and follow: Score following as a reinforcement learning game,” 2018.

[2] D. V. Balandin, R. S. Biryukov, and M. M. Kogan, “Ellipsoidal reachable sets of linear time-varying continuous and discrete systems in control and estimation problems,” Automatica, vol. 116, p. 108926, 2020.

[3] A. Baró, P. Riba, J. Calvo-Zaragoza, and A. Fornés, “From optical music recognition to handwritten music recognition: A baseline,” Pattern Recognition Letters, vol. 123, pp. 1–8, 2019.

[4] S. Hizlisoy, S. Yildirim, and Z. Tufekci, “Music emotion recognition using convolutional long short term memory deep neural networks,” Engineering Science and Technology, an International Journal, vol. 24, no. 3, pp. 760–767, 2021.

[5] Z. Wang and J. Xuan, “Intelligent fault recognition framework by using deep reinforcement learning with one dimension convolution and improved actor-critic algorithm,” Advanced Engineering Informatics, vol. 49, p. 101315, 2021.

[6] P. L. Erdős, C. Greenhill, T. R. Mezei, I. Miklós, D. Soltész, and L. Soukup, “The mixing time of switch markov chains: A unified approach,” European Journal of Combinatorics, vol. 99, p. 103421, 2022.

[6] P. L. Erdős, C. Greenhill, T. R. Mezei, I. Miklós, D. Soltész, and L. Soukup, “The mixing time of switch markov chains: A unified approach,” European Journal of Combinatorics, vol. 99, p. 103421, 2022

[7]Dorfer, M., Arzt, A., & Widmer, G. (2016). Towards Score Following in Sheet Music Images. In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR) (pp. 789–795). New York, USA.

[8]Dixon, S. (2005). An On-Line Time Warping Algorithm for Tracking Musical Performances. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI) (pp. 1727–1728). Edinburgh, UK.