week1 Jun12-18

This is the first week of my summer holiday and the beginning of my advanced project. In this week, me and my supervisor Jasper had the first online meeting and talked about my study interest as well as the structure of thesis. The start works including narrowing down research question, thinking about how to collect data, doing some case study. Jasper also told me some thing need to pay attention to, such as the research question of thesis should be focus and specific.

My research interests lies on combining machine learning as a tool into live performance. There are three missions after class, first,finding gap in knowledge. Second, search for exsiting technologies (accessible?). Third, think about technical innovation.

week2 Jun19-25

This online meeting discussed about structure of thesis. The outline of thesis usually include : 1 Abstract(to tell people what this paper about); 2 keywords; 3 Introduction; 4 Literature review; 5 Methodology; 6 Case study; 7 User study/Evaluation; 8 Discussion / Reflection.

I also narrowed my research topic which is how to build a interactive music generation tool? Instrument? which can make audience or composer interact with AI in real time. This research topic rising is because using AI to assist music creation mainly focuses on generating symbolic music melodies and tunes, and some explore interactive music creation, but there is still a lack of research in live performances where the audience can also involve the sound, so I want to study how to let the audience participate in the music In the interaction of the performance, maybe to generate some harmony? Different melodies?

week3-5 Jun26-July16

Do not have much progress. Found some papers in related field. Read the paper called *Performing with a Generative Electronic Music Controller* Link is here: https://hai-gen.github.io/2022/papers/paper-HAIGEN-MartinCharles.pdf

Week6 July17-July23

Week7 July24-July30

July 25

Today I found Ryan Kirkbride's wokr, a Python programming environment-- Foxdot and saw his video introducing Live code（Live coding concept）

Link for Foxdot is https://foxdot.org/

Link for video is https://www.youtube.com/watch?v=XRNFBZIBeuI

His personal website: https://ryan-kirkbride.github.io/

Ryan Krikbride's paper https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/eva2015.61

Collaborative interfaces for ensemble live coding performance https://etheses.whiterose.ac.uk/28901/

Also found some interesting thing:

Genetic Algorithm in python generates music, links are here

https://www.youtube.com/watch?v=aOsET8KapQQ

https://github.com/kiecodes/generate-music

https://www.youtube.com/watch?v=nypJ3b4rMhE

Using python to create systhesizer

July 26 The concept of live code:

> - Using code to describe the rules for music
> - Live notation of composition/notation as performance
> - Unlike traditional code, you can interact with with code while the program is running
> - Take computer programming into social plane

The Algorave music

July 27

Today I want to make a personal portfolio website and try this code:
https://www.kaggle.com/code/karnikakapoor/music-generation-lstm/notebook
It is a notebook shows how to use LSTM to generate music.

I found a website can build website by preset modules called squarespace.
linke here https://www.squarespace.com/

Week8 July30-Aug6

This week, I reproduced a notebook that generates music using LSTM.
The link towards this notebook is https://www.kaggle.com/code/karnikakapoor/music-generation-lstm/notebook#MODEL-BUILDING. It used dataset **Classical Music MIDI**, Source of it is here https://www.kaggle.com/datasets/soumikrakshit/classical-music-midi. The process is difficult. Chatgpt and Google were used for explaining code and solve error. The system I use is Ubuntu 22.04. I also use GPU to train model. The process of install GPU driver and set correct CUdnn and CUDA is truly difficult. I will have another part in repository to record how I solve the problem I met and the forum where I found answer.
This week I also have some new inspiration in my interaction design. Here is the background of my starting point and how the design plan forms.

BG

I want to explore the real-time interaction between humans and machine learning models in music performance. For my research project as a graduate student, I aim to achieve a small breakthrough by allowing human-generated variables to influence the music generated by machine learning models. Considering technical complexity, feasibility, and time constraints, there are two possible types of human-generated variables to be considered: one is related to changes in **human actions, such as gestures and body movements**; the other is related to physiological indicators, like **heart rate**. Based on this two possible branch, I have my concept.

Object

If **Gesuture branch** Firstly, I want to start exploring from **gestures**, focusing on how gestures can influence the generated music rhythm.
The object I want to achieve is try to train a model which can capture the change of hand （by camera） and the change of hand with lead to the change of rythm (such as position of hand, one simple way is read the coordinate of hand. If the position of hand is high, beat per minute is higher, vice versa.)
Here is a picture can explain my concept:

Then **Heart beat branch** Secondly, if the first exploration is successful, I will try build connection between music generating model with heart beat change in future.

Week8 Aug14-Aug20

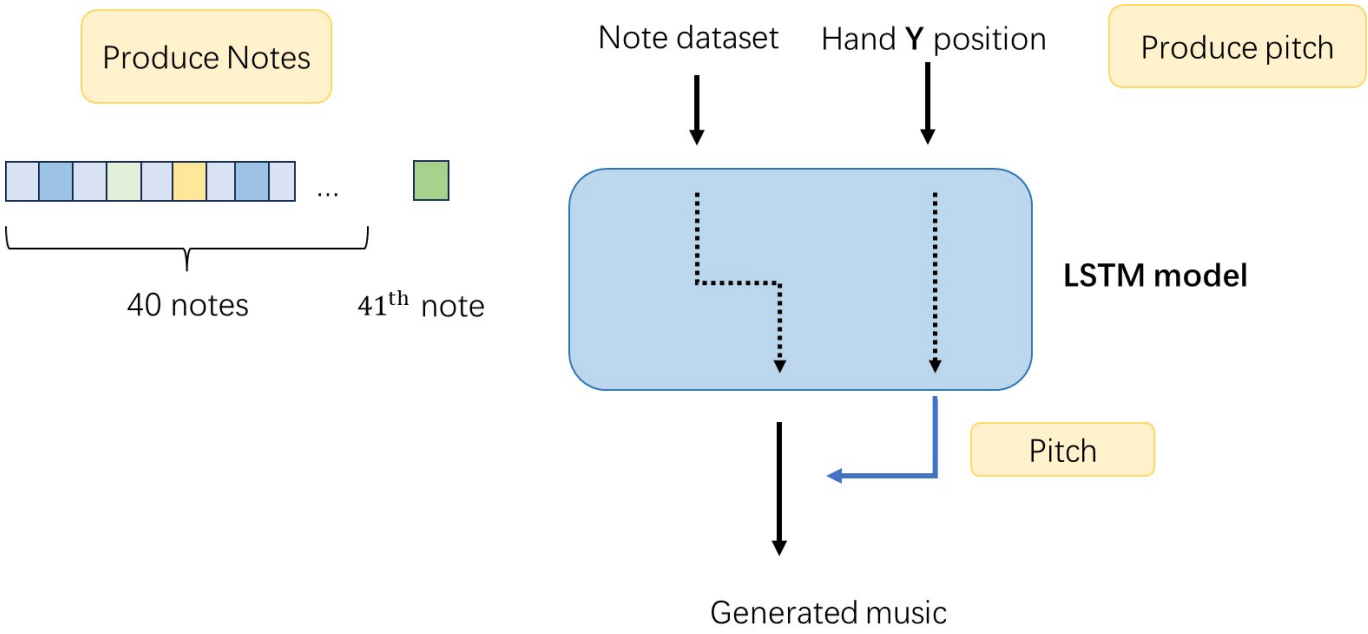
Install Openpose according to the guidance from link https://developer.aliyun.com/article/889977
Source of openpose https://github.com/CMU-Perceptual-Computing-Lab/openpose#installation Install CMake-gui for Ubuntu https://cgold.readthedocs.io/en/latest/first-step/installation.html

Aug18 Today I found Tensorflow don't have GPU version for python higher than 3.9 and music21 library require python version upper than 3.10 This is bad... And I also found Ubuntu system that I used to use doesn't support camera in my computer. Also, Ubuntu system met problem when I tried to download Openpose model.

This video really helped me to set tensorflow=2.10.0 and download Cudnn and CUDA toolkit after I have work on this problem for soooooo long time. https://www.youtube.com/watch?v=Zn6Lp0xaXj4

Week9 Aug20-Sept8

The figure below shows plan of **First generation**.



Using an example of generating notes with an LSTM as a reference, first try to reimplement this example. Then, based on the re-implemented LSTM, incorporate human creativity. The LSTM model predicts note No.41 based on the previous 40 notes. Upon capturing the Y-coordinate of the hand, it fuses the pitch corresponding to the Y-coordinate with the prior LSTM.

Achievements so far

1.Capture hand（mediapipe+camera in laptop）

2.Hand position mapping (currently only the y-coordinate has been taken)

Development of plan: The change of pitch is not obvious as duration, so I decided to start from using X coordinates to present duration,as the X coordinates changes, the duration of produced notes will change.

Store Store the document name:First_Generation

Link of Code： https://github.com/ZIqinGX/MSc_Advanced_project/tree/main/First_Generation

Evaluatioeach note or chord wn:

1.The number of X-coordinates is not as many as the dataset, so they cannot be directly integrated.

2.In the first generation of the code, as set with a fixed offset.

Week10 Oct2-Oct8

The figure below shows first solution for **Second generation**.

According to problem appeared in first generation, the solution 1 of generation 2 is using different combine strategy: LSTM model predict notes independtly then consider using X-coordinate information to adjust instead of directly make X-coordinates as input for LSTM. So the solution 1 for generation 2 is proposed.

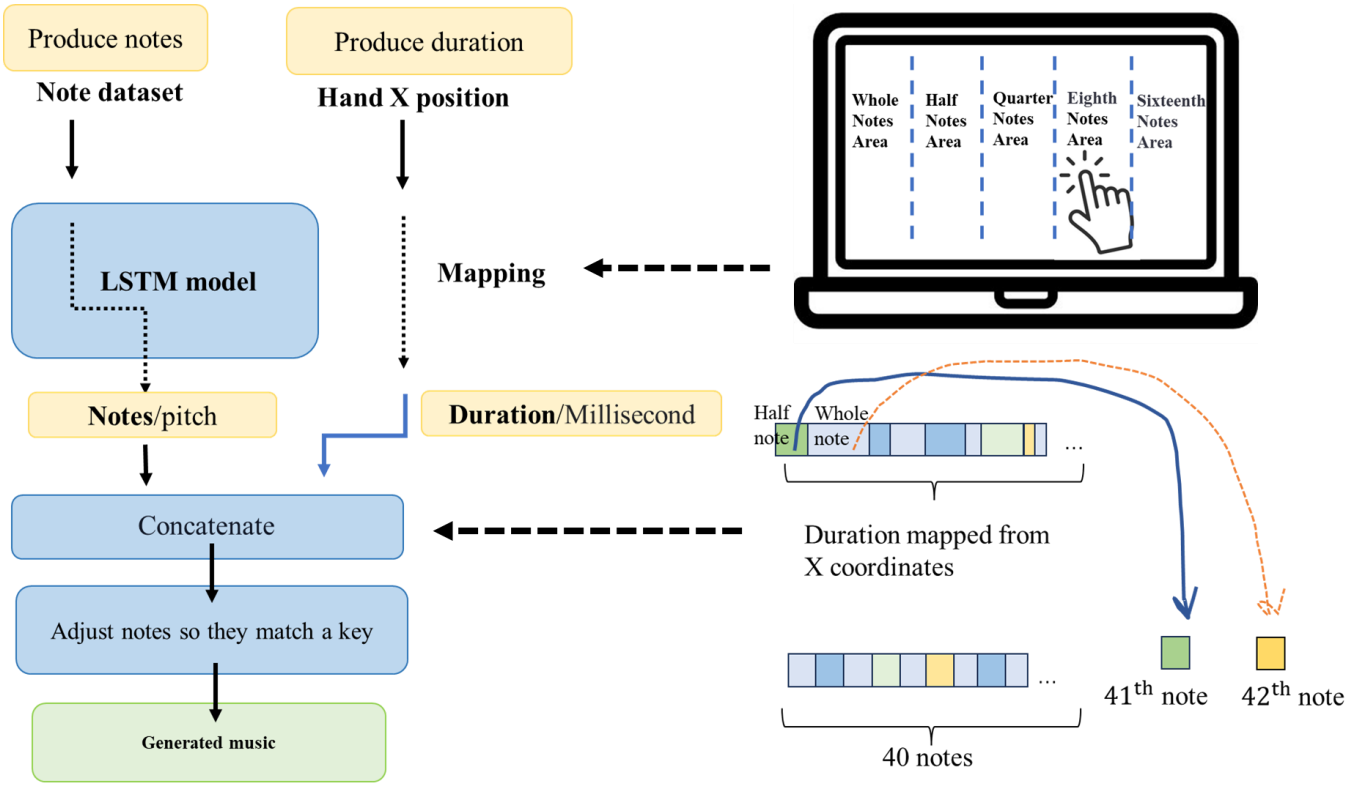
1.Mapped X-coordinates to duration to solve the lackage of coordinates information

2.X-coordinates can present duration of notes to perform meaningful tempo.

图片名称

Week11 Oct9-Oct14

This week I collect data such as velocities of hand and duration of midi file for training model in order to find out some possible connection bewteen hand movement and tempo of music. Then I found there was not previous connection between hand velocity and music tempo.



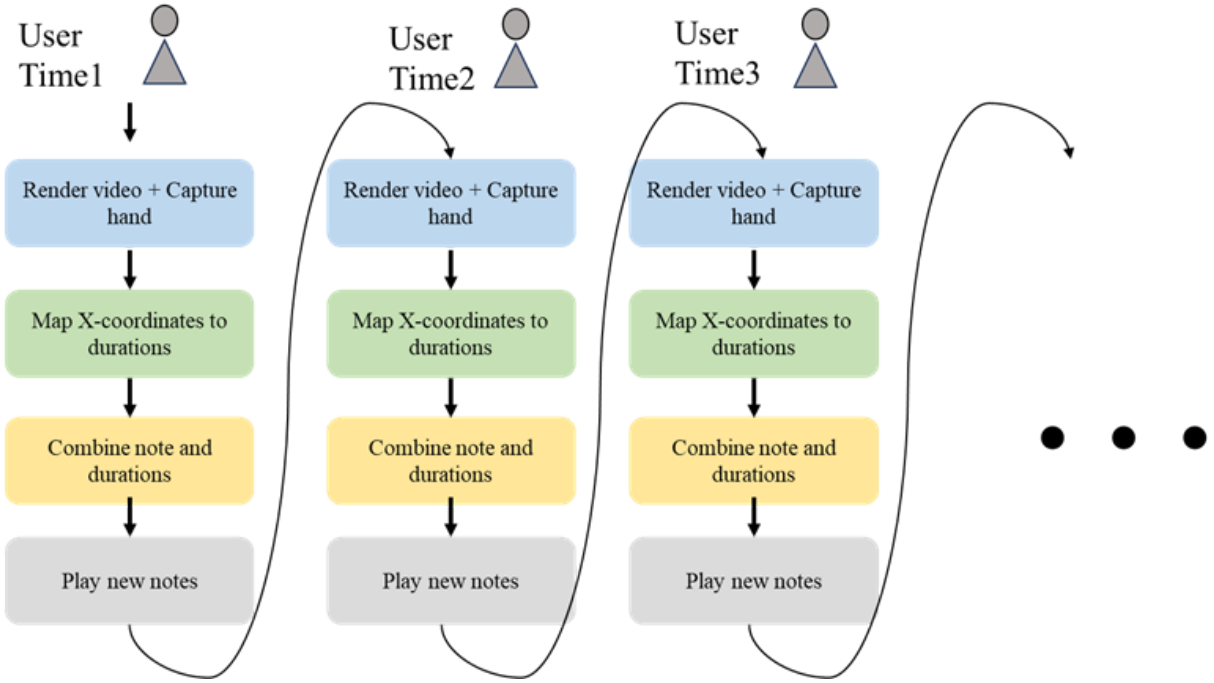
This diagram shows the development of solution2 for generation 2.

Week12 Oct16-Oct22

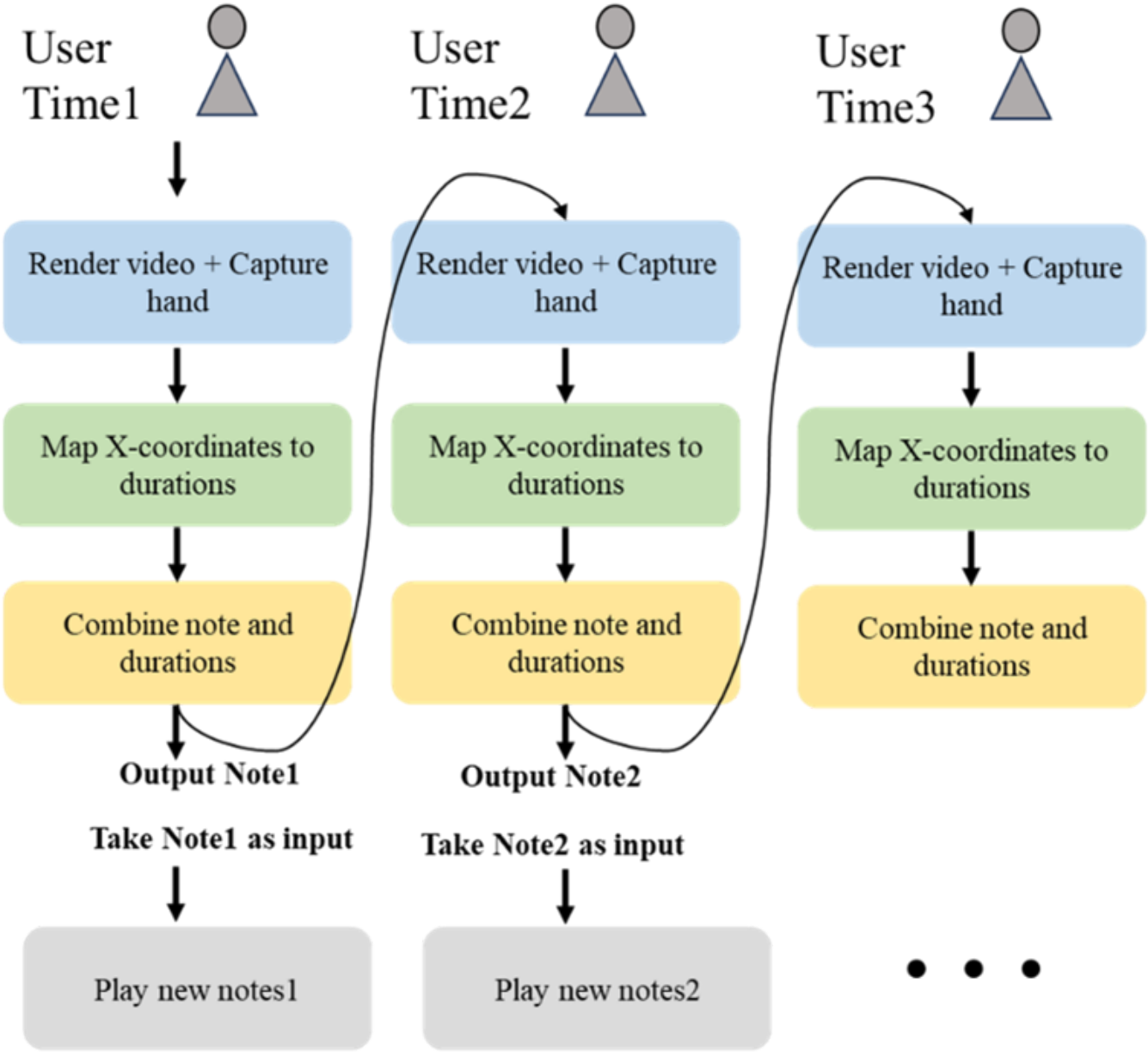
This week I make progress on realtime system. I build a parallel program to make the whole system become more senstive and respond quicker. Besides, I also trained LSTM model for more epoch.But the result shows there are still much note produced by model will be repeated. So I think maybe I can make some change in dataset because I filtered out some of the notes that occurred less frequently before I built model.

Week13 Oct23-Oct29

In this week, I found the repeated note problem can not be solved quickly, so I adopted post-processed rules to solve this problem. I also changed the serial programming of real-time system to parallel program to make user experience smoother.



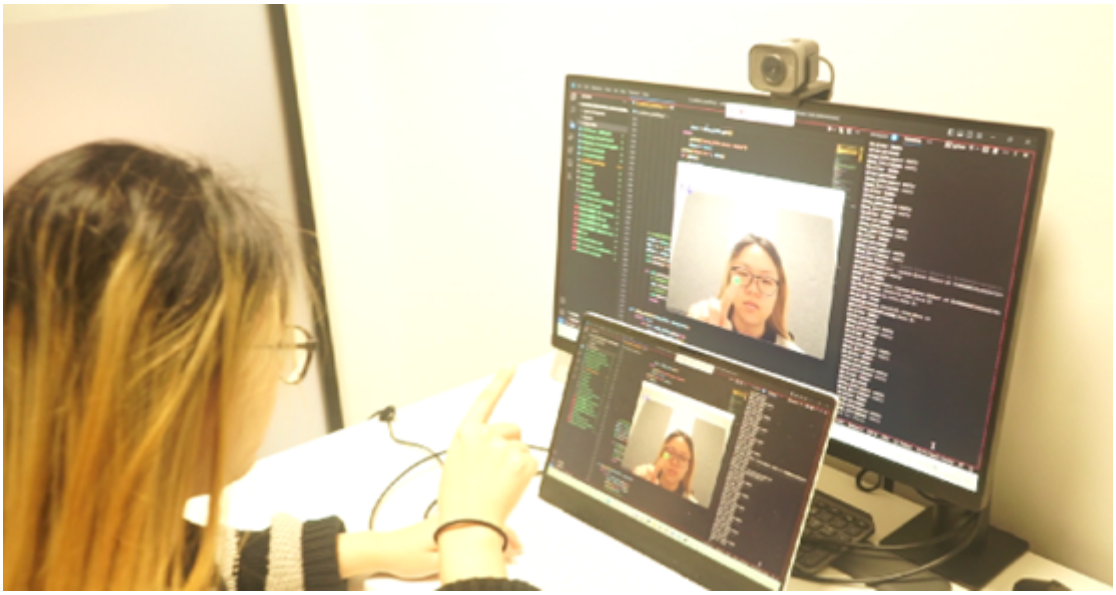
This picture is real-time system before improvement



This figure is real-time system after development

Week14,15 Oct30-Nov12

In week15, I shoot video for showing my project. I invited my model to demonstrate how to use my model. After video editing， i did the voice-over for video and keep finishing my thesis



Week16 Nov13-Nov23

In this week, I organized code, completed the paper and proof-reading. Finished weekly blog.