EIE3510 Digital Signal Processing – Progress Report
Composition and Instrumental Music Generation from Bird Songs

Yang Jiao 118010121 Jifei Zhao 118010437

26 November 2021

1. Introduction

In the AI generation, researchers started exploring the boundary between machine intelligence and artworks by using the machine to compose and generate music automatically. Previous studies mainly focused on deep learning methods, such as Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and the transformer architecture borrowed from Natural Language Processing (NLP)[1]. However, instead of the unexplainable, massive data-driven approach, we are trying to design an explicit and lightweight model with only a tiny amount of unmusical audio input.

Inspired by nature, we investigate the similarity between the singing of birds and instrumental music. With the limited amount of input audio, the generated music may have lower quality than those adopting machine learning methods. Nevertheless, to hear sounds from nature is essentially an incentive method for a musician's composition, and thus our model may also provide some insights and inspirations for musicians. In this project, a piece of recorded bird sound will be given to (1) map the sound to a series of notes (composition), and then (2) play the melody by mimicking the timbre of piano (music generation).

2. Task description and methodology

With the objective of generating music from bird sounds, this project can be divided into two stages. An integral process of this project is shown in Figure 1 in the Appendix 1. In Stage 1, our goal is dividing a piece of bird sound into segments and get the one-to-one matching of each segment with the proximate piano note[2]. For bird sounds collection, the dataset will be obtained from xeno-canto[3], which is a website contains a terrific amount of bird sounds recordings. Most of them are recorded outdoor with other background noise. Hence, it's essential to filter the noise before the subsequent operations. Short-time Fourier Transform will be adopted to obtain the spectrogram for analyzing the time-frequency property of filtered bird sound, which facilitates note division. Based on the different central frequency of segments, each audio segment will be mapped to a standard piano key with the closest key frequency. Therefore, a transition from natural bird sounds to musical note can be realized.

The next stage focuses on the melody synthesis. According to our preliminary experiment in MATLAB, the timbre of the single-tone audio output was different from the timbre of piano. To better simulate the piano sounds, harmonics will be added and the signal envelope will be properly modulated (ADSR envelope)[4]. The final step is integrating all the notes together to get a piece of melody, where cadence will be taken into consideration.

---

[1] A review on the deep learning approach can be found in this blog (retrieved on 24 November 2021): https://towardsdatascience.com/generating-music-using-deep-learning-cb5843a9d55e

[2] A table of the standard key frequencies (only including the 88 keys on a standard piano): https://en.wikipedia.org/wiki/Piano_key_frequencies

[3] Sharing bird sounds from around the world: https://www.xeno-canto.org/

[4] A blog on explaining ADSR envelope model (Chinese version only): https://zhuanlan.zhihu.com/p/62875880

Appendix 1: Task procedure

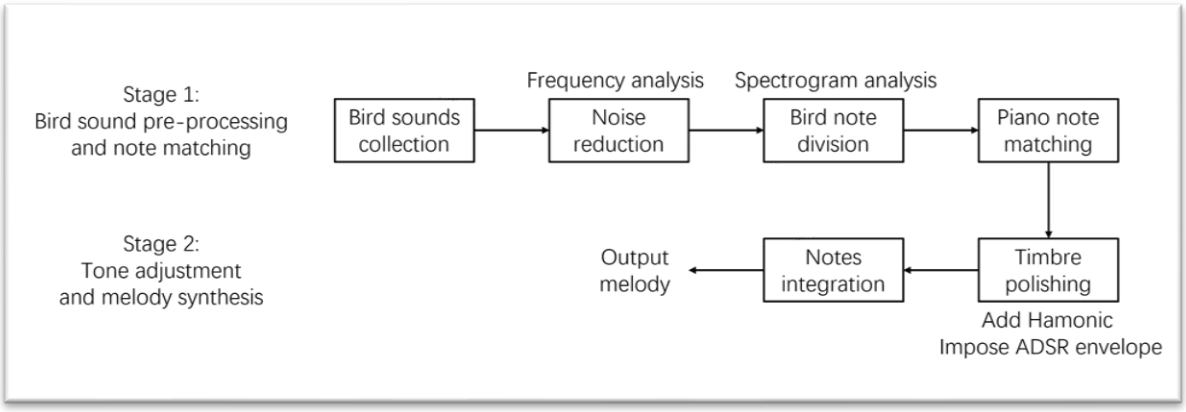

Figure 1: Procedure of composition and instrumental music generation from bird songs

Appendix 2: Work division

To implement the above model, the realization methods for each subtask will be detailly discussed in group. The work division of implementation is scheduled as follow.

Table 1: Work division among group members	
Jifei Zhao	Implementation of Stage 1.
Yang Jiao	Implementation of Stage 2; UI design for project demonstration.