

PIANO JAMMING WITHOUT A PIANO

Yichen Huang

Music X Lab, NYU Shanghai
yh2689@nyu.edu

Ziyu Wang

Music X Lab, NYU Shanghai
ziyu.wang@nyu.edu

Gus Xia

Music X Lab, NYU Shanghai
gxia@nyu.edu

ABSTRACT

We present a music piece co-created by the human performer and machine learning algorithms in an improvised fashion. Using a graphical interface, the human performer selects the chord and accompaniment texture while algorithms generate melody and piano accompaniment based on such selections. The project allows users with limited music training to create music with meaningful structures and harmony in real time.

1. CONCEPT

Whereas it is easy for one to have a general appreciation for harmony and texture in music, it takes extended practice to apply these in an improvisational setting. In this project, we explore an interactive means of improvising music for users with limited music background. The music piece is improvised in real time with the human performer selecting the desired chord and texture and machine learning algorithms generating melody and accompaniment based on these selection. In this way, the performer can control the harmony and general structure of the music without having to consider each note.

2. TECHNIQUES

As shown in figure 1, the system has three components: user interface, accompaniment generation and melody generation. The generation is conditioned on user-selected chords and reference phrases (from which texture is extracted for accompaniment). The reference phrases are a select subset of Pop909 [2], a dataset of Chinese pop songs.

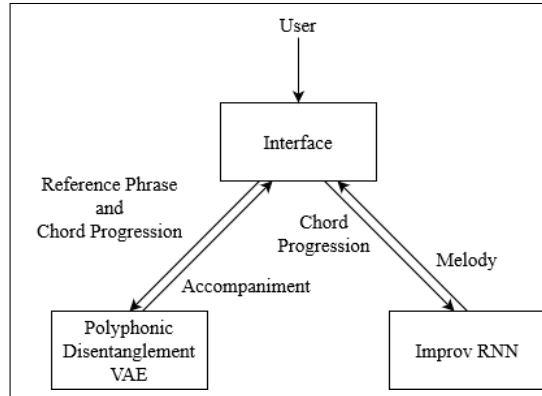


Figure 1. An overview of the system used to generate the piece.

2.1 Interface

We implement a web-based graphical interface for chord input and reference phrase selection. The interface visualizes the generated music, the current reference phrase and their respective chord progressions. The performer selects the next chord and reference phrase through a combination of mouse and keyboard inputs.

2.2 Accompaniment Generation

Accompaniment generation is achieved through harmonic style transfer using the VAE model introduced in [3]. For each two-bar window, the model generates accompaniment by combining the texture embedding of the reference phrase and the chord embedding of the input chord progression. The generation is done offline by enumerating all combinations of reference textures and chord progressions and storing the generated result. We limit the chord types and chord change frequency due to computation constraints.

2.3 Melody Generation

We generate the melody using an improv-RNN [1] conditioned on the input chord progression. Melody generation is done in real time.

3. REFERENCES

- [1] Magenta. magenta-js. <https://github.com/magenta/magenta-js>, 2021.
- [2] Ziyu Wang*, Ke Chen*, Junyan Jiang, Yiyi Zhang, Maoran Xu, Shuqi Dai, Guxian Bin, and Gus Xia. Pop909: A pop-song dataset for music arrangement generation. In *Proceedings of 21st International Conference on Music Information Retrieval, ISMIR*, 2020.
- [3] Ziyu Wang, Dingsu Wang, Yixiao Zhang, and Gus Xia. Learning interpretable representation for controllable polyphonic music generation. In *Proceedings of 21st International Conference on Music Information Retrieval, ISMIR*, 2020.