

An Automatic Music Arrangement System Using Machine Learning

Tianyi GAO[†] Hangyu SONG[†] Ahmad MOUSSA[†] Hiroshi WATANABE[†]

[†]Graduate School of Fundamental Science and Engineering, Waseda University

1. Introduction

Recently, more and more people are getting interested in music arrangement, aiming to make their own music. However, for freshman, music arrangement is too difficult to learn. An automatic music arrangement system may help this kind of people to make their music easier. With the well-developed machine learning technique nowadays, making such a system becomes possible.

2. Research Tools

2.1. Onsets and Frames

Onsets and Frames is a model for automatic polyphonic piano music transcription. Using this model, we can convert raw recordings of solo piano performances into MIDI.^[1]

2.2. Cycle-GAN

GAN (Generative Adversarial Network) is the most popular method in image generation. Inside of different kinds of GANs, Cycle-GAN performs great on image-to-image transfer, which can also be used in music style transfer.

3. Proposed Method

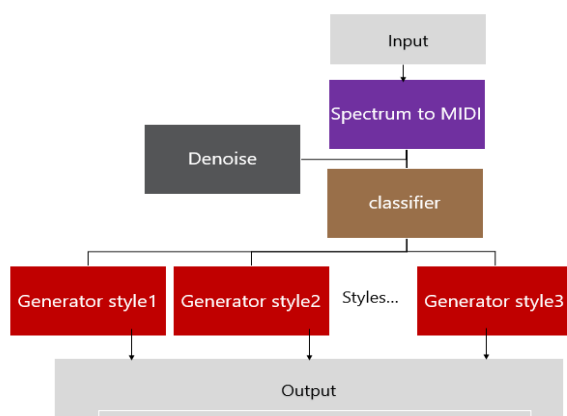


Fig. 1 Structure

The input is a melody played by piano. Using the Onsets and Frames, it becomes MIDI file, which can be understood as a music score. Then, before the generation, we made a classifier to classify which of the style fits the melody better. After that, the melody is generated to an arranged music with a pre-trained Cycle-GAN model. The output is an arranged music in MIDI file, which also can be considered as a music score.

For the Cycle-GAN model, it is supposed to be trained

with several datasets which has different styles. However, we are still making the dataset. Up to now, we have only 20 popular music available for the experiment. We catch their main melody out manually and make the melody and their arranged music in pairs. Then cut them into 16-second pieces. After that, we got 288 samples to training with. As a result, we can get the popular style for the arrangement.

4. Result

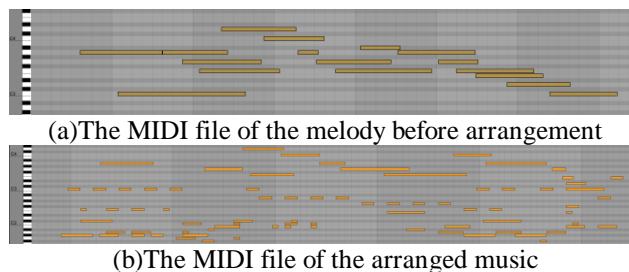


Fig. 2 Result

Fig. 2. (a) is a MIDI file of the melody we played on our piano, which is transferred from raw recordings of piano performances. Fig. 2. (b) is the MIDI file of the arranged music transferred from (a), generated by our Cycle-GAN model, trained with 288 x 16s samples of popular music. From Fig.1 and Fig.2, we can see the effect of the automatic music arrangement. We also confirmed the difference by hearing tests.

5. Conclusion

Music generation is a complex work. Different periods, different artists, even different instruments can bring us totally different styles.

Reference

- [1] Curtis Hawthorne, Erich Elsen, Jialin Song, Adam Roberts, Ian Simon, Colin Raffel, Jesse Engel, Sageev Oore, Douglas Eck “Onsets and Frames: Dual-Objective Piano Transcription”, 2018, [Online]. Available: <https://arxiv.org/abs/1710.11153>
- [2] J. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017, pp. 2242–2251.
- [3] Gino Brunner, Yuyi Wang, Roger Wattenhofer, Sumu Zhao “Symbolic Music Genre Transfer with Cycle-GAN”, 2018, [Online]. Available: <https://arxiv.org/abs/1809.07575>