AN AUTOMATIC MUSIC GENERATION METHOD BASED ON GIVEN PATTERNS OF RHYTHM

Yinji Jing
Beijing University of Posts and Telecommunications
jyj@bupt.edu.cn

Shengchen Li
Beijing University of Posts and Telecommunications
shengchen.li@bupt.edu.cn

EXTENDED ABSTRACT

This paper presents a method that generates music with given rhythm. The engagement of rhythm is fairly common for folk song composers as the lyrics can be easily transformed to the requirement of rhythm. To show the probability that the melody can be generated with the certain rhythm, the paper uses the Char-RNN (Character Recurrent Neural Network) [2] to model the relationship between the pitch line and the rhythm. The Char-RNN proposed by Andrej can generate text at the character level showing the effectiveness of RNN structure. The method proposed is based on the Char-RNN with LSTM (Long Short-Term Memory) to improve the long-time memories, the structure in this paper is shown in Figure 1. The bottom neurons is the input layer and the input is the rhythm represented as minim, crotchet, quaver, semiquaver et al. The middle hidden layer contains 1 LSTM layers with 128 units. The upper output layer output the pitch of each note represented by the absolute pitch in MIDI (which *'60'* represents *'C5'*).

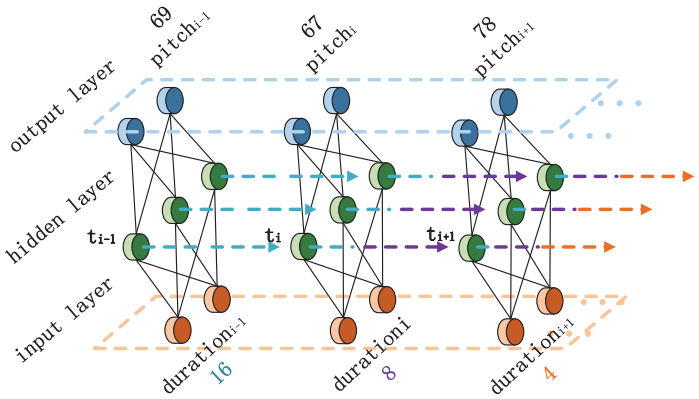

Figure 1. The RNN architecture in the paper.

In the training process, a sequence of note duration used as the input, where the cost function is cross entropy between the one-hot coded note pitches and the activation function of the output layer is softmax. In the generating process, the duration of note is input one by one, then the network gets the several most likely pitches as candidate notes. The final output is randomly chosen from the candidates for more variation in melody. The data used in this paper is MIDI format and is collected from the Nottingham Dataset which includes 1034 British folk tunes (about 200,000 notes) [1].

To get the rhythm in the generating process, the paper constructs the corresponding dictionary of International Morse Code and rhythm so that a sequence of letter can generate a certain rhythm. The dot ('•') is represented by two semiquavers ('16 16'), the dot ('–') is represented by a quaver ('8'), the space between letters is represented by a crotchet ('4'), the space between words is represented by a minim ('2'). The Figure 2 shows 6 melodies generated by the system with the rhythm of *'123456'*. The *'123456'* is firstly converted to the

Morse Code, '• – – – –, • • – – –, • • • – –, • • • • –, • • • • •, – • • • •', then the Morse Code is converted to the rhythm *'16 16 8 8 8 8 4 16 16 16 16 8 8 8 4 16 16 16 16 16 16 8 8 4 16 16 16 16 16 16 16 16 8 4 16 16 16 16 16 16 16 16 16 4 8 16 16 16 16 16 16 16 16'*.

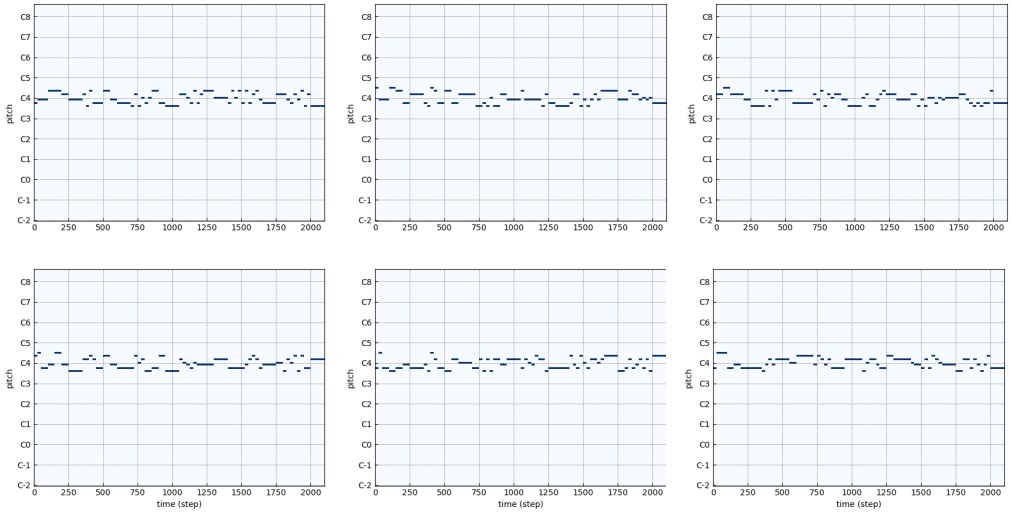

Figure 2. Melodies with the same rhythm pattern.

To evaluate the quality of generated music objectively, the log-likelihood of 50 music, which is the degree of similarity between the music Nottingham Dataset and the music generated by the system, is calculated. The MEAN_LLH is the mean value of log-likelihood, the STD_LLH is the standard deviation of log-likelihood. The MELODY SOURCE is the source of music compared to the dataset, the *Nottingham Dataset* refers the music composed by human and the candidate notes is the most likely pitches to choose. The Table 1 shows that the log-likelihood of each piece of music composed by human ranges from -2.547 to -3.341 covers the log-likelihood of music generated by the system when the candidate notes are less than 5. The results shows that the melody generated by the given rhythm can be similar to the human composed melody and also shows the potential of the RNN in generating music information with music structure like rhythm pattern.

MELODY SOURCE	MEAN_LLH	STD_LLH
Nottingham Dataset	-2.959	0.412
1 candidate note	-2.880	0.013
2 candidate notes	-3.087	0.051
3 candidate notes	-2.931	0.076
4 candidate notes	-3.275	0.100
5 candidate notes	-3.420	0.134

Table 1. Log-likelihood of fifty music compared with dataset

ACKNOWLEDGMENTS

The authors would like to thank to the Zhou Wei team of ZHONGWEN LAW FIRM for funding this project.

REFERENCES

[1] Jukedeck. Cleaned version of the nottingham dataset. `https://github.com/jukedeck/nottingham-datase`. Mar 3, 2017.

[2] Andrej Karpathy. The unreasonable effectiveness of recurrent neural networks. *Andrej Karpathy blog*, 21, 2015.