Proceedings of The 2021 IEEE International
Conference on Real-time Computing and Robotics
July 15-19, 2021, Xining, China

RE-RLTuner: A topic-based music generation method

Haibin Liu[1,2], Xurong Xie[1,3], Rukiye Ruzi[1], Lan Wang[1,4], Nan Yan[1,4]

Abstract— **Automatic music generation has become a more exciting research topic. However, the existing music generation methods tend to utilize music data inherently and hardly consider music generation from the perspective of music composition theory. Therefore, how to use music theory to guide the automatic generation of music is drawing increasing attention to this area. In this work, we aim to extract music rules from given corpora and then apply them to generate new music of a similar style. We divided the melody into different scales (music segments with different numbers of notes, not set of notes ordered by pitch) based on the music structures and then employed Latent Dirichlet Allocation (LDA) topic model to learn the structural constraints of the provided musical form. The multi-scale fusion of musical features through reinforcement learning (RL) enables the model to consider the music generation from the global scope. Our experimental results show that our model is superior to the baseline model according to objective and subjective rating. The music generated from our model has better consistency in terms of music style, which indicates that the extracted structural features and the multi-scale modeling are promising for music generation of a certain style or topic.**

I. INTRODUCTION

Music is an essential part of our life as it is a way of expressing our feelings and emotions. In the last decades, digitalization promoted both accessibility and demand for music. To meet the increasing needs of music products, researchers both from computer science and digital music technology attempt to realize computational music generation technology. Recently, deep learning and modern generative modeling techniques have been implemented in the automatic generation of music. Unlike rule-based methods[12] which require manually extracted features, deep learning methods could avoid complex rule design and explore higher-dimensional features of music[14]. Generative Adversarial Network (GAN)[2], [3], [4] has also been applied to music generation, where random noises serve as an input to the generator whose goal is to transform random noises into the objective. GAN-based methods often achieve outstanding results by carefully designing the generating strategy of the discriminator. Yet, it requires a large amount of data to train the network, making the training process quite challenging. As Long Short-Term Memory (LSTM) network architectures excel in modeling sequential information in data, LSTM based music generation method is proposed in[7]. It predicts the following note by constantly repeating the previous note's input, where gives promising results. Although subsequent work has optimized the LSTM[8], it suffers both from the mismatch between training and testing conditions and the mismatch between training criterion and generation objectives in music generation[11].

The current technology often pursues music data[2], [3], [4] and rarely considers it from the music composition perspective. Music has distinctive structural characteristics. In music generation, it is necessary to start from the musical structure[1], considering the kinds of music typically prevalent in public are very well-structured. Therefore, to model music structure characteristics and applying them to music generation is of great significance.

To apply music theory to music generation, a novel sequence learning approach, RLTuner was proposed in [12]. In this model, Reinforcement Learning is used to impose structure on an RNN. Task-related rewards and the probability of a given action, which is learned from the pre-trained RNN, are combined by the reward function in RLTuner. This model is the first application of reinforcement learning in music generation. However, the list of rules could not be exhausted. RL-Duet[11] was also put forward as a real-time accompaniment music generation system, which captured the context of music through a bidirectional LSTM on the reward function. This algorithm is designed to generate a melodic, harmonic and diverse machine part of the music response to the human part. Even though subjective evaluations on preferences show that the proposed algorithm generates music pieces of higher quality than the RLTuner method, the main disadvantage of data-driven reward models is that it only focuses on the superficial contextual relationships of the notes, but can not make the most of the structural characteristics of the music [1].

Topic Model is widely used in data mining, a recommendation system, and some other fields[13], [10]. Researchers define the same commonness of data as a topic[16]. Latent Dirichlet Allocation (LDA) is one of the most successful algorithms for the topic model. The LDA model is based on a definite theorem to capture significant inter- and intra-document statistical structure via mixing distributional assumes that documents arise from multiple topics. Since a topic is a distribution over document, LDA topic model can find the data?s internal connection in the massive data. At the same time, music of the same style has similar characteristic features in different scale[1]. Thus, we aspire to explore whether the music sequence also has the same theme utilizing the LDA method in topic features extraction.

In this study, we introduce RE-RLTuner, a multi-scale

[1]CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.
[2]University of Chinese Academy of Science, Beijing 10049, China
[3]Chinese University of Hong Kong, Hong Kong 999077, China
[4]Guangdong-Hong Kong-Macao Joint Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen 51800, China
*Corresponding author:Nan Yan(e-mail: nan.yan@siat.ac.cn)

2021 IEEE International Conference on Real-time Computing and Robotics (RCAR) | 978-1-6654-3678-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/RCAR52367.2021.9517538

modeling method for music generation which makes use of music structure characteristics. Since the topic model is used to model the music at different scales and the reward function can capture the structural information of music, the generated notes are the product of multiple aspects. Our main contributions are as follows: a novel music feature extraction method based on the LDA topic model is proposed. And adopting the multi-scale music information to guide music generation, we present the RE-RLTuner, which generates music pieces of higher preferences than the baseline method concerning the integration of music characteristics at different scales.

II. Methodology

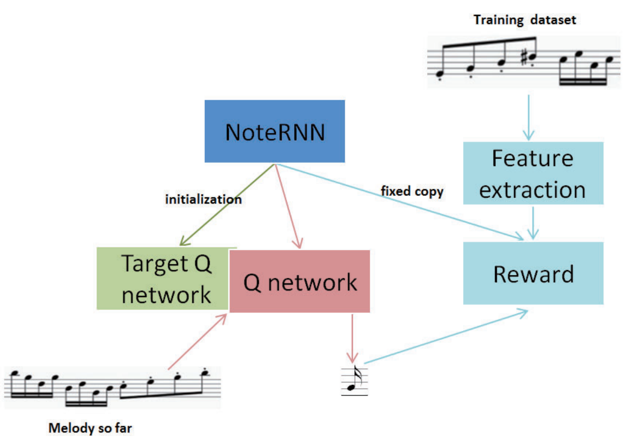

Fig. 1: Model struct, It includes music feature extraction, reward function and reinforcement learning

Figure 1 illustrates the overall structure of our model. We use the deep Q learning(DQN) reinforcement learning architecture. Q network and target Q network are DQN training tricks where the target Q network is a delayed copy of the Q network. Feature extraction is the LDA-based topic feature extractor and Reward is the Multi-scale modeling design.

A. Music feature extraction

In this section, the LDA model is used to cluster music data sets at different scales to model the music structure features. Gibbs sampling method is adopted to calculate the LDA model. One component of the joint distribution is sampled each time, and the other components are kept unchanged. After derivation, Gibbs is used to solving the LDA topic model, and its calculation formula is as follows:

$$p(z_i = k|z_{\bar{i}}, w) = \frac{n_{k,\bar{i}}^{(t)} + \beta_t}{[\sum_{z=1}^{K} n_k^{(v)} + \beta_v] - 1} \times \frac{n_{m,\bar{i}}^{(k)} + \alpha_k}{[\sum_{z=1}^{K} n_m^{(z)} + \alpha_z] - 1} \quad (1)$$

In Eq.1, $\bar{i}$ means deleting the i-th item, $n_k^{(v)}$ means the number of times the music segment v appears in the k topic. β_v represents the Dirichlet prior distribution of the music segment v. $n_m^{(z)}$ means the number of times the topic z appears in the m document. α_k is the Dirichlet prior distribution of topic v. $p(z_i = k|z_{\bar{i}}, w)$ is the conditional probability distribution of a certain music segment ω_i corresponding to the topic feature z_i. The conditional probability of Gibbs sampling is used to sample the topics. After multiple iterations converge, the topic of music segment distribution β and the document topic distribution θ can be obtained. This music structure feature extractor utilizes the LDA topic model algorithm and takes music segments as input to get the corresponding topic model. We extract topics at different scales and define them as topic1, topic2, topic3.

B. Music reward

At present, the existing reward function mainly can be divided into two types, manual rules-based reward function[12] and context-information-based reward function[11]. Manual rules-based reward functions tend to be crude and their application is confined to a particular set of data[11]. The context information-based reward functions take only a limited amount of information into account. In this paper, we design the reward function for different scales of given corpora, that is to get the $p(note|Topic)$.

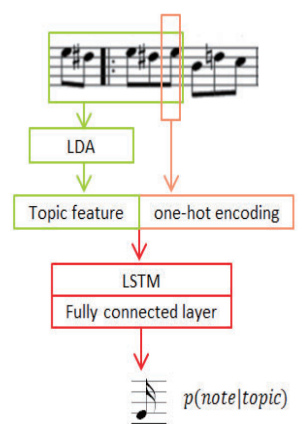

Fig. 2: topic-based model structure

We convert the note sequence into a single hot encoding and concatenate it with the current topic feature obtained from the previous music structure feature extractor. Then, the concatenated feature is further used as input for the reward model. Our reward model consists of LSTM connected with a fully-connection layer. A note and its probability are generated by the reward model based on previous music note sequences and topics. The probability serves as the reward model?s actual reward value. All three topic models comprise a similar structure. Different music structure information is integrated through topic weight, thus completing the music reward function model. Finally, we get three music generation models which can carry different music structure information.

C. RL

The traditional supervised learning method only considers the short-term return. In contrast, the reinforcement

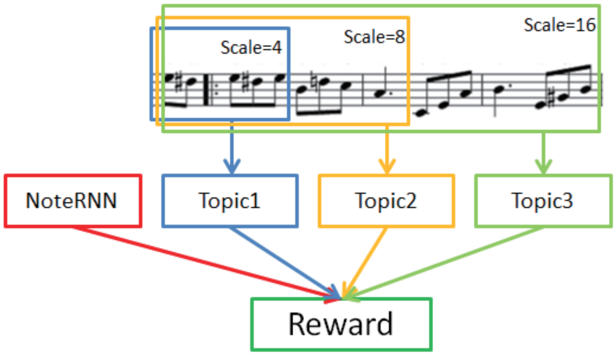

Fig. 3: Reward design

learning method, as a solution to the sequential decision problem, aims at maximizing the long-term return. Inspired by RLTuner, we treated the music generation problem as a sequential decision problem using the classical DQN algorithm. Figure 1 shows the framework of our model. We integrate music structure features at different scales through reinforcement learning.

NoteRNN is trained as a single-layer LSTM network, consisting of Q network, Target Q network and RewardRNN. Target Q network is the delayed copy of the Q network. RewardRNN is fixed together with the reward function to maintain its characteristics. We take the state of the current Q network and the generated music sequence as the state s of the agent, a as the selected note, and r as the reward function of reinforcement learning. $p(a|s)$ generated based on the probability of action for Q network ω_i respectively for each topic corresponding weights, $p(note|topic_i)$ for the corresponding topic generated on the probability of the note. γ, θ, $Q(s, a; \theta)$ represent learning rate of reinforcement learning, current strategy of reinforcement learning and a selection under the current policy respectively. The reward for the note at time t is:

$$r_t = \ln p(a|s) + \sum\nolimits_i^{topic} \omega_i \ln p(note|topic_i) \qquad (2)$$

Eq. 3 is the loss function of DQN and Eq. 4 is the policy of DQN.

$$L(\theta) = \mathbf{E}_\beta[(r_t + \gamma \max Q(s^{'}, a^{'}; \theta^-) - Q(s, a; \theta^-))^2] \quad (3)$$

$$\pi_\theta = \arg\max_a (Q(s, a; \theta)) \qquad (4)$$

III. Model Detail

A. Dataset

In the field of neural network based music generation, music data format can be divided into audio and symbol representations. While a few works are carried out in audio for its cumputational demanding, most works are focused on the symbol domain. In our experiment, we focus on the symbol domain. Carrying the duration and pitch information of music, symbolic representation has simpler computation characteristics and therefor it is popular among researchers[12]. We use the Nottingham dataset, a dataset of folk songs, pitch range from 53 to 88, which we quantified into 4/4 beats.

B. Model Specification

Feature extractor: Since we need to model music at different scales, it is necessary to conduct experiments on the classification of music segments in different units. Empirically, we chose a number of notes as a music unit, which represents the span of the scale. There are 4, 8 and 16 notes in the scale of topic1, topic2 and topic3 respectively. And those three topics consist of 20, 12 and 5 categories correspondingly.

Reward function: Due to the amount of parameter, we set the LSTM hidden state to 15 and stitched a full connection layer to predict the notes. The accuracy of the verification set was 76%.

RL: In the RL module, NoteRNN's structure is similar to the reward function for it only takes the music sequence as input. In the training of RL, the learning rate of the Q network is 0.01, the discount rate of reward is 0.95, and the update rate of the target Q network is 0.01. As for the weight of the topics, we set 0.2,0.3, and 0.1, respectively.

C. Generation

Considering the nature of the LSTM model, using maximum likelihood estimation, repeat too many times can lead to the high probability of generated notes. To solve this problem, we sampled the generated music notes. Our model produces notes and their corresponding probability, where the probability is regarded as a polynomial distribution. Additionally, we randomly choose note sequence to set the start bar for the training process.

IV. Experiments

In this section, we compared the performance of RLTuner, LSTM, and Re-RLTuner models. The LSTM we used is NoteRNN. The reward function of RLTuner is composed of a series of music modules, including interval control, auto-correlation coefficient, and the relationship between scales and intervals. In our implementation, we used the default parameters in RLTuner[12].

A. Objective Evaluation

To verify the validity of the model, we adopted the following indexes for quantitative analysis as in [5], [12].

	PC/bar	PI	IOI	Auto-lag1	Auto-lag2	Auto-lag3
dataset	1.51	25.7	1.49	0.88	0.77	0.67
LSTM	2.73	20.0	13	-0.41	0.44	-0.37
RLTuner	3.6	28.02	18	-0.03	-0.03	-0.03
RE-RLTuner	2.9	17.5	**11**	**−0.10**	0.003	**−0.01**

PC/bar represents the number of different notes in each bar. PI represents the maximum interval of the bar. IOI represents the average interval of the bar. Auto-lag represents the correlation of the music sequences, and N represents the distance of several units. We found that the improved method is better than RLTuner in PI, IOI, and correlation indexes compared to those of the dataset. Intuitively, notes from RLTuner are too dense and messy e.g. PC/bar value is 3.6, of the highest, while notes from RE-RLTuner show a good correlation with each other and have reasonable intervals.

B. Subjective Evaluation

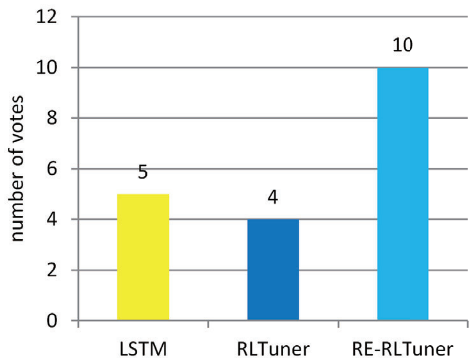

Fig. 4: Subjective Evaluation, the number of volunteers prefer the generated music

To obtain more intuitive results, we randomly recruited 19 volunteers as audiences. We randomly chose four music segments from each of the LSTM, RLTuner, Re-RLTuner, and a piece of original music from the dataset. We asked volunteers to listen to the original music from the dataset and based on that to rate extracted music samples individually. In addition, they were also asked to evaluate how close each of the three types of music samples to the music from the dataset. The score ranges from 1 to 5. We found that

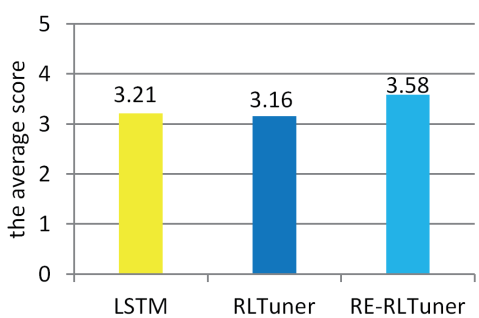

Fig. 5: Subjective Evaluation, the volunteers? ratings of the similarity in style of the generated music and the original music.

RLTuner is close to the LSTM score, but RE-RLTUNER outperforms the baseline by subjective score. In terms of style consistency, the music generated by RE-RLTuner has a clear lead. Besides, RE-RLtuner shows a clearer and more natural style, which indicates that our model is capable of capturing music style with the topic model.

C. discussion

In our experiment, we extracted music structure features through the LDA topic model and achieved the multi-scale fusion of music features through reinforcement learning. Because LSTM simply captures the context aspect of the music, just a limited part of the learned music is repeated in the generated music. The performance of RLTuner is not appealing and we assume that it is because the parameters cannot be well controlled due to the complexity of manual rules. The structural features of the music extracted through the LDA topic model are effective in keeping the style consistent, which is verified in the evaluation. Our model is exceeding the baseline model in music style.

V. Conclusion And Future Work

In this work, we extracted music rules to guide the generation in order to get a similar style. We dealt with the data at different scales and then employed LDA topic model to learn the structural constraints. Reinforcement learning is implemented to the multi-scale fusion of musical features for addressing the music generation from the global scope. However, due to the limitation of LSTM's own long time series, our future work can be carried out on reinforcement learning agents.

VI. acknowledgement

This work was supported by the National Key R&D Program of China [grant number: 2020YFC2004100]; the National Natural Science Foundation of China [grant numbers: NSFCU1736202, NSFC61771461]; the Shenzhen Fundamental Research Program [grant numbers: JCYJ20170413161611534], and the Science and Technology Planning Project of Guangdong Province [grant number: 2019B090915002].

References

[1] Newlin D , Schoenberg A , Strang G . Fundamentals of Musical Composition[J]. Music Educators Journal, 1968, 55(1):92.

[2] Mogren O . C-RNN-GAN: Continuous recurrent neural networks with adversarial training[J]. 2016.

[3] Dong H W , Hsiao W Y , Yang L C , et al. MuseGAN: Symbolic-domain Music Generation and Accompaniment with Multi-track Sequential Generative Adversarial Networks[J]. 2017.

[4] Yu Y , Canales S . Conditional LSTM-GAN for Melody Generation from Lyrics[J]. 2019.

[5] L. Lerch. On the evaluation of generative models in music[J]. Neural Computing and Applications, 2018(2).

[6] Hasselt H V , Guez A , Silver D . Deep Reinforcement Learning with Double Q-learning[J]. Computer ence, 2015.

[7] Eck D , Schmidhuber J . Finding temporal structure in music: Blues improvisation with LSTM recurrent networks[C]// IEEE Workshop on Neural Networks for Signal Processing. IEEE, 2002.

[8] Sturm B, Santos J F, Korshunova I. Folk music style modelling by recurrent neural networks with long short term memory units[C]//16th International Society for Music Information Retrieval Conference. 2015.

[9] Roberts A, Engel J, Hawthorne C, et al. Interactive musical improvisation with Magenta[C]//Proc. NIPS. 2016.

[10] Ramage D, Hall D, Nallapati R, et al. Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora[C]//Proceedings of the 2009 conference on empirical methods in natural language processing. 2009: 248-256.

[11] Jiang N, Jin S, Duan Z, et al. Rl-duet: Online music accompaniment generation using deep reinforcement learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(01): 710-718.

[12] Jaques N, Gu S, Turner R E, et al. Generating music by fine-tuning recurrent neural networks with reinforcement learning[J]. 2016.

[13] Wang Z Z , Ming H E , Yong-Ping D U . Text Similarity Computing Based on Topic Model LDA[J]. Computer Science, 2013.

[14] Briot J P , Pachet F . Music Generation by Deep Learning - Challenges and Directions[J]. Neural Computing and Applications, 2020, 32(2).

[15] Briot J P , Hadjeres G , Pachet F . Deep Learning Techniques for Music Generation - A Survey[J]. 2017.

[16] Chen Q , Yao L , Jie Y . Short text classification based on LDA topic model[C]// 2016 International Conference on Audio, Language and Image Processing (ICALIP). IEEE, 2017.