MusicCollage: A Music Composition Tool for Children Based on Synesthesia and a Genetic Algorithm

Ge Yan[1], Cheng Yao[2(✉)], Chao Zhang[1], Jiadi Wang[1], Yuqi Hu[3], and Fangtian Ying[4]

[1] College of Software Technology, Zhejiang University, Hangzhou, China
{yange,zhangchaodesign,wjd1024}@zju.edu.cn
[2] College of Computer Science and Technology, Zhejiang University, Hangzhou, China
yaoch@zju.edu.cn
[3] International School of Design, Zhejiang University, Ningbo, China
yuqihu@nit.zju.edu.cn
[4] College of Design, Hubei University of Technology, Wuhan, Hubei, China

Abstract. Creating music can be challenging for children aged 4–6. Making the creation procedure more interesting has been proved to stimulate and build children's interests in music. Based on the design concept of synesthesia from vision to hearing, we propose an application by which children aged 4–6 can create collages by matching different images of scenes and animals. The information contained in the created collages is then used as input parameters for generating music so as to create melodies corresponding to the collages. To evaluate our approach, we conducted a quantitative and qualitative user study with children (n = 6) by using our initial prototype. We concluded that the proposed application helps cultivate children's willingness to create melodies, and helps children better perceive and understand music by combining scenes and animal images in the collages.

Keywords: Synesthesia · Music creation · Children · Collage · Genetic algorithm

1 Introduction

The benefits of a quality music education are immense. Interaction with music is important for the development of young children. Exposure to music from childhood onward can help children speak more clearly, develop a larger vocabulary, and strengthen their social and emotional skills [1]. Musical skills can also give people, both children and adults, joy during their lives. There are also indications that musical development helps children with their overall development [2]. This is why an effort must be made to try to stimulate children's interest in music composition.

Music creation requires a certain knowledge of music theory as a background, which is a great challenge for children. However, schools and parents lack the appropriate music education for children. Music education in many public school jurisdictions has been downgraded to insignificant or non-existent because of the expense [3]. In addition,

N. A. Streitz and S. Konomi (Eds.): HCII 2022, LNCS 13326, pp. 205–216, 2022.
https://doi.org/10.1007/978-3-031-05431-0_15

children are less likely to be encouraged in this direction if their parents did not benefit from a quality music education. For these reasons, we established a design principle to simplify music-making to make it easier for children to participate in music creation and enjoyment.

There have been previous studies on designing children's music creation tools. LEGO Bricks [4] was found to provide a physical and social interaction environment for young children. Amal Tidjani [5] designed MuSme, a tangible skin suit that uses different parts of the body as metaphors for different musical instruments. The interactive software designed by Yin Ling [6] takes synesthesia into account, which refers to the establishment of the correspondence of color and shape (e.g., circle, line, and square) to a music scale and note, but this refers to a color or shape corresponding to a single note or scale, without the correspondence of the whole picture. It is difficult to complete a melody by only relying on a combination of notes for children.

With the help of synesthesia, children can better perceive music. In addition, the technology of artificial intelligence (AI) lowers the threshold of creation, making the process easier [7]. According to this idea, we developed a web application based on Brower/Server (B/S) architecture. Children can use the application by accessing a browser. Children can create a complete picture by selecting different scenes, matching different kinds of animals, and adjusting the placement of those animals. The information in the collage will then be used as input parameters of the music generation algorithm to ultimately create unique melodies based on a genetic algorithm.

Before using the application, one of the researchers gave the children a brief introduction to it and a demonstration of the process. Six children aged 4–6 completed at least one collage with this application in 20 min. After confirming that the collage did not need to be modified, the corresponding melody could be played. At the end of the experiment, we asked the children to submit the questionnaires and describe their feelings about using the application. Finally, we talked to parents who had observed the whole experiment and recorded their thoughts.

The main contributions of this paper are:

- We designed a synesthesia-based strategy by matching different scenes and animals to cultivate the interest of 4–6-year-old children in creating music and improve their perception of music.
- We developed a web application based on B/S Architecture, whereby children can create their own pictures and corresponding unique music by selecting different scenes and animals and adjusting the placement of animals.
- Using quantitative and qualitative methods, we conducted user studies to verify the effectiveness of using this application to cultivate children's interest in music creation and improve their ability to perceive music.

2 Related Work

2.1 Children's Music Education and Creation

Musical education has a positive impact on children's intelligence, creativity, and memory [8, 9]. Many studies on musical education for children have been conducted to help

them grow and develop. Ian McKinnon [3] outlined the development of an interactive software program called "Children's Music Journey" to deliver affordable and pedagogically rigorous music education to children everywhere, and as a result, improve attitudes toward learning and promote better behavior in school. Leyi Ouyang's research [10] proved that tangible designs have great potential in engaging young children in early music learning concepts, such as rhythm, pitch, dynamics, and expression; they provide an environment for young children to interact physically and socially, which in turn arouses their curiosity and sense of fantasy. Steve Altieri [11] produced an interactive tool called "keyboard" for students and gives them an opportunity to "play around with music." Yin Ling [6] designed an interactive visual music education software that integrates scales, intervals, and harmonies. The product enables children to increase their musical knowledge and develop their musical perception and understanding. These rich forms of interaction stimulate children's interest and make music easily accessible to them.

The creative process enables children to better understand and learn about music. Children composing music without a threshold can make them appreciate the joy of music more. TempoString [12] is an easy-to-use tool that assists children with music creation. It provides a fun and novel platform by allowing children to "draw" music on a canvas. Amal Tidjani [5] proposed MuSme, a tangible skin suit that reimagines a user's limbs and organs as metaphoric representations of different instruments. Jane Rigler [13] designed Cre8tor, an interactive music composition system controlled by motion sensors that allows children to create music while playing. Uwe Oestermeier [4] presented a multi-touch tabletop application that utilizes LEGO bricks as physical representations for musical notes to create a novel digital learning environment for musical composition principles. PlaceAndPlay [14] presented a new design of a digital music authoring tool for children. It incorporates multimodal interaction techniques, including voice-based synthesis of instruments and sound effects, to assist younger users who do not play musical instruments or do not have knowledge of sound effects. The above research demonstrated the wonderful integration between children and music creation, and it inspired us to design a platform for children that is fun, accessible, and allows them to create music while playing.

2.2 Synesthesia of Sound and Image

People's senses are closely interconnected, and visual and auditory experiences are often associated with each other. Many researchers have conducted interesting research on the synesthesia of sound and image. Current research has indicated that the application of synesthesia can effectively improve the user experience. Michael Voong's research [15] indicated that there is a reasonable degree of consistency between users' associations of color and music, and that an indirect descriptor can aid in the recall of music via mood. Siyu Jin [16] designed an interactive work where users select music and images as the main interactive contents, and the parameters of the music are used as the dynamic expression of human emotions. The new pixel generation process of the image is regarded as the result of emotions affecting humans. Vegas [17] explored the interactive potential through live performances and the space for aesthetic expression by synthesizing the audio and the visuals. This project related to the genre of visual music and abstraction

in the arts and created a synesthetic experience for the audience. Dimitris Kritikos [18] developed a deterministic process to produce a melody after processing a painting, mimicking the production of notes from colors in the field of view of persons experiencing synesthesia.

Motivated by synesthesia design, we imagined that correlating colors and sounds would enable children to create images in a form that intuitively generates music of the same emotion.

2.3 AI Music Creation

AI has great potential in the field of music creation, made possible by advances in machine learning and algorithms that allow people without expertise to lower the threshold of creation and be inspired with the assistance of AI [7]. Emma Frid [19] presented a user interface paradigm that allows users to input a song to an AI engine and then interactively regenerate and mix AI-generated music. This solution makes music generation easily and understandable. Ryan Louie [20] found that AI-steering tools not only enabled users to better express musical intent, but they also had an important effect on users' creative ownership and self-efficacy vis-a-vis AI. Cong Ning [21] designed a musical tabletop application that enables both novice and musically trained users to compose and play music in a creative and intuitive way. This tangible musical interface supports finger operations and aims to lower the barriers to composing music in order to enrich the music-playing experience. These studies demonstrated the power of AI in the field of music creation. We hope that with the assistance of AI, the creative process for children will become easier, and the content they create will be richer.

3 MusicCollage

3.1 User Flow

The whole process of interacting with MusicCollage is divided into the following steps. First, children choose their favorite scenes, which is important, as it establishes the style of the whole melody. Then, children can choose the animals they want to appear in the collage and can drag the animal images in the selector to the positions. During this process, the positions of the images can be adjusted. To delete an animal that was dragged onto the screen last time, children simply need to click the undo button on the screen. Children can also click the reset button if they want to remove all of the animals. After that, they can continue adding animals. Finally, after confirming the scene and animals, children can play the melody created with the picture information as the input. Both the collages and the generated melodies will be saved on the server (Fig. 1).

3.2 System Design

MusicCollage consists of two main parts: one is a picture generation module, and the other is a music generation module.

Picture generation module: The homepage of the application contains a scene selector, an animal selector, and a canvas (Fig. 2). By clicking the start button in the center of

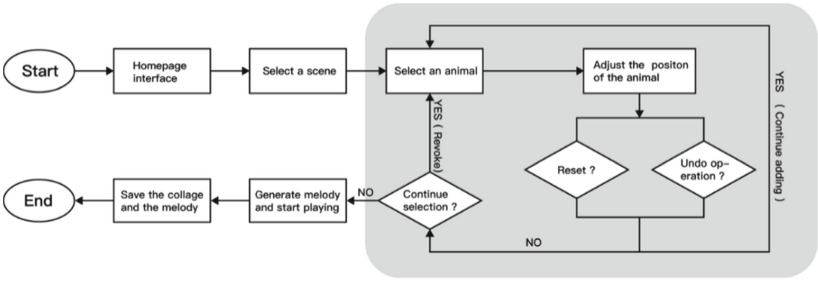

Fig. 1. Workflow of MusicCollage

the canvas, the child can begin the journey of creating the collage. Afterward, children can start using a scene selector and an animal selector. The background of the canvas is temporarily set to the first background of the scene selector by default. Children can drag the images of the scene selector horizontally to view various scenes. The scene selector can be dragged horizontally to view various scenes. After clicking on the desired scene, the background of the canvas will be replaced by the selected scene. The animal selector allows children to view animal images by dragging them vertically. After choosing the animal, children can click on and drag the desired animal into the screen. The scene selector contains eight scenes, and different scenes have different picture styles. Similarly, the animal selector includes 17 animals, including elephants, hippos, raccoons, monkeys, and rabbits. As shown in Fig. 3, children can click the reset button or undo button to complete the removal of all of the animals or the last one, respectively. If the child is satisfied with the collage, they can enjoy the music after clicking the play button.

Fig. 2. The homepage interface includes the (a) canvas, (b) scene selector, (c) animal selector, and (d) start button.

Fig. 3. The user interface includes the (a) undo button, (b) reset button, and (c) play button.

Music generation module: We considered the following two goals: 1. Ensure consistency and correlation between the information in the picture and the input parameters in the music generation algorithm. Consistency refers to the exact same picture, and the corresponding algorithm input parameters should also be consistent. Correlation means that the correspondence between the picture information and input parameters should be reasonable and in line with people's understanding and cognition. 2. We should focus on cultivating children's interest in creating music so that the melody created by children is unique. Specifically, even identical drawings with exactly the same input parameters should not produce exactly the same music. For these reasons, we chose to use an open-source genetic-algorithm-based music generation algorithm [22]. It contains a total of 10 input parameters (Table 1).

Table 1. Input parameter of the algorithm

Name	Description
Number of bars	Length of the generated melody in bars
Notes per bar	Number of notes inside of a bar
Number of steps	Number of pitches per note
Introduce pauses	Should the algorithm introduce pauses between notes or do you want a constant stream of notes?
Key	Key of the scale the melody should fit in
Scale	Type of scale the melody should fit in
Scale root	Pitch of the scale (ex.: 4 means C4 is the root note of a scale in C)
Population size	Number of melodies per generation to rate and recombine
Number of mutations	Max number of mutations that should be possible per child generated
Mutation probability	Probability that a mutation will occur

3.3 Synesthesia in Design

We used the total number of animals as the parameter Number of bars. Then, we used the distance between the leftmost horizontal axis coordinate and the rightmost horizontal axis coordinate divided by the value of the total number of animals to map the parameter Notes per bar. Specifically, the more animals there are, or the smaller the distance between the animals in the left-most and right-most positions, the smaller the parameter. Next, we used the mean value of the vertical axis coordinate center to map the parameter Number of steps. As for the parameter Key, we represented it by the rounded average value of the corresponding value of animals (Table 2). For example, when the average value of all of the animals is 9.6, the input parameter Key of the algorithm will be Gb. How these numbers correspond depends on whether the animal is more likely to be perceived as deep and big or cheerful and flexible. Parameter Scale is one of the following: major, minorM, Dorian, Phrygian, Lydian, mixolydian, major blues, and minorBlues. Each type of input corresponds to a scene (Fig. 4). As for the parameter Introduce pauses, we chose the default parameter true as suggested by the algorithm because it is more in line with the rules of music creation. We also chose the default parameter 4 when coming to the parameter Scale Root. Finally, when it came to the input parameters of the genetic algorithm itself, including Population Size, Number of mutations, and Mutation probability, we still selected the default parameter after trying several input parameters.

Table 2. Value of parameter key and animal

Value	Key	Animal	Value	Key	Animal
1	C	elephant	10	Gb	sheep
2	C#	hippopotamus	11	G	fox
3	Db	crocodile	12	G#	dog
4	D	bear	13	Ab	raccoon
5	D#	lion	14	A	hedgehog
6	Eb	tiger	15	A#	monkey
7	E	panda	16	Bb	rabbit
8	F	giraffe	17	B	mouse
9	F#	donkey			

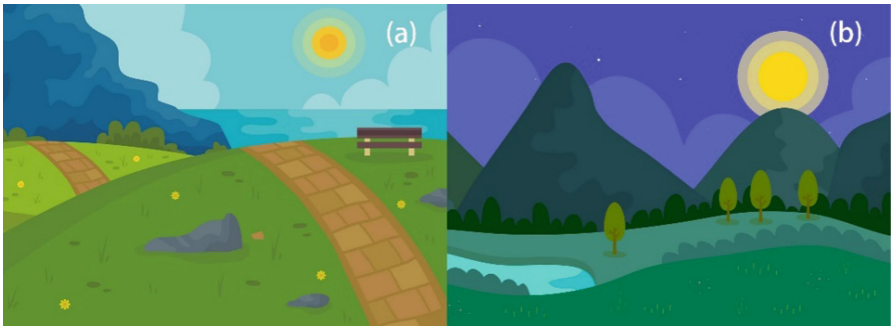

Fig. 4. (a) The scene that corresponds to "major;" (b) the scene that corresponds to "minorBlues."

4 Evaluation

4.1 Participants

We recruited six children to participate in our experiment. They ranged in age from 4 to 6 years (M = 5.0, SD = 0.58). Three were boys, and three were girls. Six parents were also invited to observe our experiment.

4.2 Procedure

Before the experiment, a letter was sent to the children's parents informing them of the study and asking for their consent.

Children were given a 10-min introduction to the application's functions and general structure, followed by a five-minute demonstration of how the system works, including how to build pictures and how to play music.

They were then given 20 min to create their own favorite drawing using the scene selector and animal selector. They could change scenes at will, add or remove animals, and adjust the placement of animals. After the final confirmation of the picture, the music was played (Fig. 5). In the whole process, children could create more than one picture, resulting in more than one piece of music.

After the experiment, the children were asked to fill out a questionnaire for a quantitative study. Each child was then asked to describe how they felt about using the application. We also conducted semi-structured interviews of about 10 min each with parents who had observed the process.

Each child participating in the experiment and each parent interviewed received a gift worth $10 for participating. In addition, the researchers gave the children pictures and music generated by the server.

4.3 Measurements

The questionnaire for children consisted of the four questions that follow. We used a Likert scale and asked the children to rate each question on a scale of one to five. Five means strongly agree while one means strongly disagree.

Q1: Do you think it is fun to write melodies this way?

Q2: Would you like to use this application often to create melodies?

Q3: Is the feeling you get from the pictures similar to the feeling you get from the melodies?

Q4: Can you use the application to create the melody you want to hear?

The questions asked during the semi-structured interviews were as follows:

- Do you think our system helps children understand music?
- What do you think can be improved about our application?
- Would you be willing to let your child use it for a long time?

Fig. 5. Children interacting with MusicCollage in the experiment.

5 Result

5.1 Quantitative Analysis

Q1 and Q2 from the questionnaire are related to cultivating children's interest in music creation, and Q3 and Q4 are related to enhancing children's musical perception. Through the feedback of the questionnaire, we found that children still showed high interest in using this application. Moreover, the children showed a high degree of agreement that different music could be created through the application.

In the end, we obtained a moderate score on the question of whether the application could create the music children wanted. This may have stemmed from the fact that some children do not fully understand synesthesia from sight to hearing. Another reason is that there is room for improvement and iteration in the algorithms that generate music (Fig. 6).

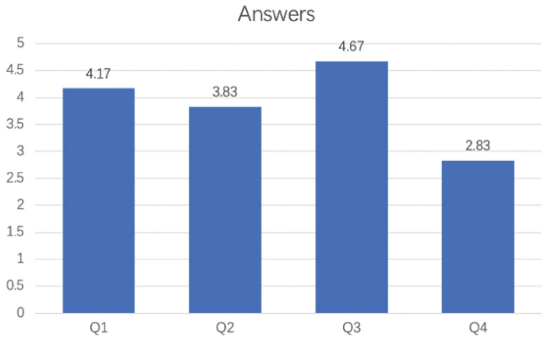

Fig. 6. Average of the answers children gave for the following questions after interacting with MusicCollage: Q1: Do you think it is fun to write melodies this way? Q2: Would you like to use this application often to create melodies? Q3: Is the feeling you get from the pictures similar to the feeling you get from the melodies? Q4: Can you use the application to create the melody you want to hear? (measured by Likert scale on a scale of one to five and five means strongly agree while one means strongly disagree)

5.2 Qualitative Analysis

Children generally believed that the application was interesting. Four children tried to create music many times during the experiment, and C2 even created it 12 times. C1, C2, C3, and C5 told us that they liked the pictures they created and were also satisfied with the music they created. It is worth mentioning that C5 told us that he would like his kindergarten teachers to experience the application.

In the interview with parents, P1 mentioned that her child was obviously excited when using the application. P2 told us that he thought the application was interesting, and felt the effect of the generated music was better than he expected. P3 asked whether children could be allowed to freely doodle on the picture instead of simply choosing the picture. What attracted our attention was that P4 was unwilling to let his child use this application for a long time because he did not want his child to touch the electronic screen, believing it might affect the child's eyesight. P5 proposed trying to use it herself and finally tried to create three pieces of music, and expressed satisfaction with two of them.

We made an innovative attempt to design an engaging application for children to create music by drawing pictures based on the design concept of synesthesia. Most of the children showed great interest in using the application. On the whole, children were satisfied with the pictures and music they had created, and some children showed a strong interest in the application and repeatedly created music. In our communication with the parents, we found that the parents as a whole also thought highly of the application. Some parents pointed out that they liked the design method based on synesthesia and could fully understand it, but they were worried that some children could not fully understand the change from picture style to music style. The most amazing thing for us was that one parent proposed to try using the application themselves. After creating pictures and music, she expressed that it was a good experience.

6 Limitations and Future Work

There are still some areas that can be improved in the application. For example, the types of elements can be more diversified rather than limited to scenes and animals, and children can be allowed to doodle to a certain extent on the picture so as to accommodate children's tendency to doodle. Some parents worried that this application would affect their children's vision, and we hope that this problem can be solved.

We will continue our research in the future. First, we will compare more open-source music generation algorithms. At the same time, we will try to improve the algorithm by including chords, remixes, and other effects to perfect the sound. In addition, we will try to add some new elements in the application, such as clouds or trees, to enrich the content of the picture and enable the whole picture to contain more information. Also, the application could provide more input parameters for the generation algorithm. We also plan to add different instruments so that children can hear the effects of their favorite instruments, which will make the system more lively and interesting. Finally, we plan to complete a combination of physical methods, such as replacing the scene selector and picture selector in our system with canvas and hard cards, to present a similar effect; this will avoid eye injury caused by electronic screens to children.

7 Conclusion

In summary, we designed and created a music creation application for children aged 4–6 that generates music through drawing; it is conducive to cultivating children's interest in music creation and enhancing their musical perception. In contrast to the previous systems, we aimed to stimulate children's interest in actively making music. A fun format was designed to make it easier for children to create and understand music.

Overall, this study is considered to allow children to experience the joy of music creation while drawing, and provide children with a visual and auditory experience at the same time. We designed and developed MusicCollage to bring the experience of music composition to children and improve their enthusiasm for music creation.

Acknowledgments. This research was funded by the Engineering Research Center of Computer Aided Product Innovation Design, Ministry of Education, Fundamental Research Funds for the Central Universities, National Natural Science Foundation of China (Grant No. 52075478), and National Social Science Foundation of China (Grant No. 21AZD056).

References

1. Ostrov, J.M., Gentile, D.A., Mullins, A.D.: Evaluating the effect of educational media exposure on aggression in early childhood. J. Appl. Dev. Psychol. **34**(1), 38–44 (2013)
2. Taylor, H.: Music with the Under-Fours by Susan Young. Routledge Falmer, London and New York (2003). (139 p. £ 14.99, paperback. Br. J. Music Educ. **21**(1), 131 (2004))
3. McKinnon, I.: Children's music journey: the development of an interactive software solution for early childhood music education. Comput. Entertain. **3**(4), 1–10 (2005)

4. Oestermeier, U., et al.: LEGO music: learning composition with bricks. In: Proceedings of the 14th International Conference on Interaction Design and Children, pp. 283–286. Association for Computing Machinery, Boston, Massachusetts (2015)
5. Tidjani, A., Cho, E., Lee, P.: MuSme: a tangible skin suit for music creation. In: Proceedings of the TEI 2016: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, pp. 743–748. Association for Computing Machinery, Eindhoven, Netherlands (2016)
6. Ling, Y.: Interactive visual music education software for children. In: Proceedings of the 20th International Academic Mindtrek Conference, pp. 453–456. Association for Computing Machinery, Tampere, Finland (2016)
7. Moruzzi, C.: Creative AI: music composition programs as an extension of the composer's mind. In: Müller, V.C. (ed.) PT-AI 2017. SAPERE, vol. 44, pp. 69–72. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-96448-5_8
8. Brown, L.L.: The benefits of music education. PBS KIDS for Parents (2012)
9. Jaschke, A.C., et al.: Music education and its effect on intellectual abilities in children: a systematic review. Rev. Neurosci. **24**(6), 665–675 (2013)
10. Ouyang, L., et al.: Designing tangible interactions with children for pre-school music education. In: 32nd Australian Conference on Human-Computer Interaction, pp. 687–691. Association for Computing Machinery Sydney, NSW, Australia (2020)
11. Altieri, S., et al.: Music education technology for older elementary school aged children: playing around with mini Amadeus Jr. J. Comput. Sci. Coll. **20**(5), 102–103 (2005)
12. He, L., et al.: TempoString: a tangible tool for children's music creation. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 643–644. Association for Computing Machinery, Pittsburgh, Pennsylvania (2012)
13. Rigler, J., Seldess, Z.: The music creator: an interactive system for musical exploration and education. In: Proceedings of the 7th International Conference on New Interfaces for Musical Expression, pp. 415–416. Association for Computing Machinery, New York (2007)
14. Akiyama, Y., Oore, S.: Place and play: a digital tool for children to create and record music. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 735–738. Association for Computing Machinery, Florence, Italy (2008)
15. Voong, M., Beale, R.: Music organisation using colour synaesthesia. In: CHI 2007 Extended Abstracts on Human Factors in Computing Systems, pp. 1869–1874. Association for Computing Machinery San Jose, CA, USA (2007)
16. Jin, S., Qin, J., Li, W.: Draw portraits by music: a music based image style transformation. In: Proceedings of the 28th ACM International Conference on Multimedia, pp. 4399–4400. Association for Computing Machinery, Seattle, WA, USA (2020)
17. Šimbelis, V.V., Lundström, A.: Synthesis in the audiovisual. In: Proceedings of the CHI 2016 Conference Extended Abstracts on Human Factors in Computing Systems, pp. 301–304. Association for Computing Machinery, San Jose, California, USA (2016)
18. Kritikos, D., Karpouzis, K.: From pixels to notes: a computational implementation of synaesthesia for cultural artefacts. arXiv preprint: arXiv:2101.12038 (2021)
19. Frid, E., Gomes, C., Jin, Z.: Music creation by example. In: Proceedings of the 2020 CHI 2020 Conference on Human Factors in Computing Systems, pp. 1–13. Association for Computing Machinery, Honolulu, HI, USA (2020)
20. Louie, R., et al.: Novice-AI music co-creation via AI-steering tools for deep generative models. In: Proceedings of the CHI 2020 Conference on Human Factors in Computing Systems, pp. 1–13. Association for Computing Machinery, Honolulu, HI, USA (2020)
21. Ning, C., Zhou, S.: The music pattern: a creative tabletop music creation platform. Comput. Entertain. **8**(2), 13 (2011)
22. Kunigk, D.: Generate-music (2020). https://github.com/kiecodes/generate-music