*Human Systems Engineering and Design (IHSED2023),* Vol. 112, 2023, 63–72
https://doi.org/10.54941/ahfe1004114

# DJaytopia: A Hybrid Intelligent DJ Co-Remixing System

**Yue Wu[1], Anran Qiu[1], Liuxuan Ruan[1], Xuejie Li[1], Jinhao Huang[1], and Stephen Jia Wang[1,2]**

[1]Intelligent Systems Design Specialism, Master of Design, School of Design, the Hongkong Polytechnic university, Hongkong SAR, China
[2]Laboratory for Artificial Intelligence in Design, Hong Kong Science Park, New Territories, Hong Kong Special Administrative Region, China

## ABSTRACT

Nowadays, musical mixing platforms are featured with programmed interventions and digitized information visualization to support DJ's performance (Montano, 2010), however, the visualization is always obscure to the average music consumers (Beamish, Maclean, and Fels, 2004). Being a well-performed DJ requires the level of expertise and experience that most average music consumers lack (Cliff, 2000), as every audience has a completely different taste in music (Schäfer and Sedlmeier, 2010). This study aims at developing an AI / ML-based system to lower the bar for novice DJs and even average music consumers to create personalized music remixes. Generally, music can be intelligently composed by analyzing harmonic and melodic features to generate genre-specific compositional elements or to alter the compositional structure of a song (Tan and Li, 2021). Despite the technical breakthroughs that have been made, listeners have reacted negatively to this music due to the lack of user data to back it up and the neglect of the user's perception of the piece (Tigre Moura and Maw, 2021). In a conventional scenario, DJs can express their attitudes towards music preferences by listening to the music directly, which requires a well understanding of the audience's mind. Following the recent launch and explosion of ChatGPT, which has evidenced that an intelligent system could help users innovate by solving their problems in textual form through conversational interactions (Dis et al., 2023; Dwivedi et al., 2023); also collecting the users' feedback through conversations, observing user reactions, and inviting user reviews. Such AI-enabled systems are able to learn about the user's preferred music style and various DJ mixing techniques. This study adopts a typical human-in-the-loop (HITL) approach to develop a crowd-learning music mixing system implementing AI and Virtual Reality technologies. The proposed HITL-based co-music arrangement system should be able to collect musical data and techniques; a VR environment is built to provide users with a platform to record user-created music and corresponding applied methods as well as audience ratings worldwide. After processing the data, users can try out a compilation of songs assisted by a robotic arm. With the help of the robotic arm, it will be easier and faster for users to create collections with a personal touch and more specific techniques. The essential functions include: a) Providing users with an immersive environment to learn the basic operations of the DJ console. b) Collecting the user's preferences for compilation techniques and the content of different DJ's compositions for use through an "immersive online multiplayer music compilation platform" to generate a personalized library of methods to help the user compile songs; c) Assisting the user in creating their preferred individual compilation style faster as they try out the DJ's operations; d) Indicating to the user where the music needs to be equalized, switched or arranged. Instead of showing the user the digital music signal to assist in creating more efficiently, the system directly operates on the DJ console. User experience experiments were conducted with both novice DJs and experienced DJs to validate whether the proposed system could help humans in creating more engaging music with stronger musicality. Five participants, respectively three novice DJs and two experienced DJs, joined two experiments of half an hour on a virtual DJ and an actual DJ console. They started the experiment by experiencing the virtual DJ console and DJ community in VR. They remixed independently first and then collaborated with the robotic arm together for music production on the actual DJ console. Three different audience also joined the experiment to evaluate the performance of users. The result was that the music produced with the robotic arm had better musicality. The user's attitude towards the whole experience, reflected in whether the music was rhythmic or the system was inspiring was recorded in the feedback. Overall, the users had a satisfying and smooth experience, and the collaborative music remixing had a certain level of musicality, but there is still some room for improvement in terms of user understanding. However, the users expressed that this fresh collaborative approach made them more interested in DJing and motivated their desire to learn and create.

**Keywords:** Human-robot collaboration, DJ music creation, Machine learning, Virtual reality

 

## INTRODUCTION TO CURRENT DJ MUSIC REMIXING

Musical mixing platforms now employ various features such as programmed interventions and digitized information visualization to help DJs in their performances (Montano, 2010). These features display aspects of the music such as beat synchronization, parameter analysis, and content context (music genre) analysis to assist users in making more informed choices during the music production process (Schwarz and Fourer, 2021). However, such visualizations can be obscure to the average music consumer (Beamish, Maclean, and Fels, 2004). Without basic knowledge of music, such as rhythm and beats, average consumers may find it challenging to understand the principles of DJ performances. On the other hand, even a well-performed DJ cannot meet the satisfaction of every audience as they have completely different taste in music (Schäfer and Sedlmeier, 2010). These factors pose a significant barrier to becoming a good DJ, especially for those who create great remixes that resonate and interact with the audience.

Generally, music can be intelligently composed by analyzing harmonic and melodic features to generate genre-specific compositional elements or to alter the compositional structure of a song (Tan and Li, 2021). For DJ music production, existing algorithms can generate music accompaniment segments using machines as DJs, including tagging music features, automatic transitions between songs, and adjustments to high, medium, and low-frequency music, all of which are done using digital music signals on computers (Callander, 2022). Although this facilitates DJs' work to some extent, it lacks the performative aspect. In addition, DJ performances also involve the audience's reactions and atmosphere, and music produced through computers often cannot receive positive feedback from the audience (Tigre Moura and Maw, 2021; Huang et al., 2022). In a conventional scenario, DJs can express their attitudes towards music preferences by listening to the music directly, requiring a good understanding of the audience's mind. This requires DJs to adjust the content of the songs according to the audience's attitudes during the performance. This is where audience feedback and DJ performances need to be combined, which is based on user feedback. ChatGPT is a model based on user conversational feedback.

Following the recent launch and explosion of ChatGPT, which has evidenced that an intelligent system could help users innovate by solving their problems in textual form through conversational interactions (Dis et al., 2023; Dwivedi et al., 2023); also collecting the users' feedback through conversations, observing user reactions, and inviting user reviews. Such AI-enabled systems are able to learn about the user's preferred music style and various DJ mixing techniques, which can help improve the richness of the performance and capture more user preferences.

Thus, our system aims to improve DJ music creation by integrating the users' feedback then replacing the traditional method of audio file generation, to achieve "collaborative remixing" with robotic arms.

## METHODOLOGY

To build the entire human-robot co-creation system (HR-CCS), human-in-the-loop (HITL) method was applied in this system. HITL has been verified

to be applicable in the context of machine learning. HITL can change the data in the machine learning model to adapt to user needs rather than random sampling (Wu et al., 2022; Kumar, Kaur, and Singh, 2020). In this system, to address the contradiction between audience feedback on music and DJ style, and to bring new styles to remixing music, HITL can help improve our machine learning model. To enable the robotic arm to better understand user needs and gain user trust for collaborative DJ music mixing, the entire system will first collect user preferences on music styles and mixing techniques. To provide users with multiple mixing patterns and styles and obtain more diverse data, a VR system in a virtual environment is constructed to collect such data. These data will be processed and computed, such as by machine learning, and finally used on the robotic arm to serve collaborative music production between users and the robotic arm. Based on this relationship, users play a crucial role in the entire system. Initially, they provide data, and then the data provided by users is used for themselves. Therefore, the relationship between the system and users is bidirectional. The system can adapt and adjust according to user needs, and users can also adjust and optimize the collaborative co-creation process based on the machine's feedback.

## SYSTEM DESIGN IN ROBOTIC ARM SUPPORTED DJ MUSIC CO-CREATION WITH HUMAN

A skilled and experienced DJ has the ability to execute a diverse range of performing actions, including looping, scratching, mixing, and comprehending EQ settings. With the number of effects at a DJ's fingertips, it can be intimidating or confusing for new DJs to know which effects they should learn and how to use them. Furthermore, even experienced DJs can face challenges utilizing these effects to their full potential. It requires a great deal of manual dexterity and coordination to manipulate multiple knobs and buttons simultaneously. For new DJs, this is where the robotic arm can come in to provide a solution.

In our design, the robotic arm controls the Pads of sound effects in the DJ controller, making the DJ's performance stand out and leaving a lasting impression on the audience. Sound effects, also called Effects (or FX) has been considered as one of the most critical essential tools for DJs. Effects change the characteristics of the sound by changing the preset options of the track, it can react to natural phenomena (like reverb, echo, etc.), or it can be a digital signal, simulating some artificial effects, adding color and depth to the track. The range of effects available to a DJ encompasses a broad spectrum, spanning from slight modifications to the sound to outright distortion. For example, a DJ might use a filter effect to gradually reduce the high frequencies of a track and create a build-up to engage the audience, or they use a delay effect to create an echo of a sound to add texture to a track. These effects provide multiple solutions for creating music tracks and making smooth transitions between different songs. Also, the usage of Effects is not only as the most basic music creation tool but also as a mode for masters to show off their flexibility in handling sound.

So in concert with the robotic arm, new DJs do not have to be afraid of complex operations and sound effects selection and can freely control knobs and faders, and other intuitive and more accessible parts since the system can help them learn DJ skills more quickly and collaboratively create music with the robotic arm as a partner, rather than relying on traditional computer output music files. Professional DJs can feel the joy of creating music with the robotic arm, according to the effect given by the robotic arm partners to make a rich and exciting collision. As a reliable partner, the users can master more new styles as well as create ground-breaking music with the system.

## DEVELOPING AN ROBOTIC ARM SUPPORTED DJ MUSIC CREATION SYSTEM

To implement this system, a virtual DJ environment was built on Oculus Quest. In this environment, users can experience the operation of a DJ console (See Figure 1). The sides of the DJ console control different song playback and effects, simulating the functionality of a real-life DJ console. Users are allowed to select songs from the menu above and play with tricks on the DJ console to create their own remixing. Based on the special properties of the VR environment, DJ performances that were once limited to a specific location can now be experienced by music enthusiasts around the world through VR glasses in an immersive way.


**Figure 1:** Virtual DJ console for users to learn and operate in VR.

'Data on user preferences has been collected and this forms the basis for iterative enhancement of the model for machine learning using the HITL approach in VR, where more targeted samples are used to improve the accuracy of the model. In this system we are supported by Edge Impulse. Data styles from VR are first manually filtered into drums and uploaded to Edge Impulse with labels to set up the database. Features were then generated and models were built with the appropriate sample size, algorithm, etc. (see Figure 2) selected. We finally achieved an idealised model and downloaded it locally.

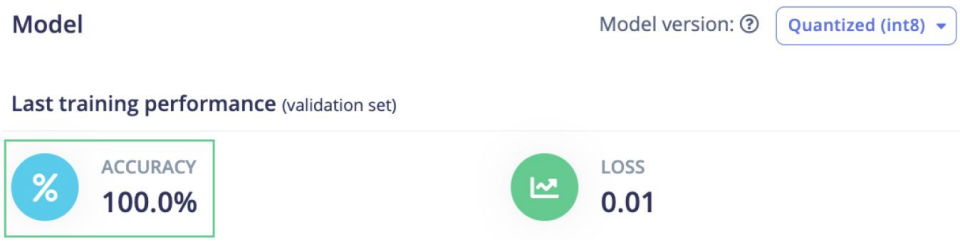


**Figure 2:** Machine learning model in edge impulse.

In order to apply the model to the robotic arm and allow it to respond to the real-time audio, the Arduino Uno was used as a bridge between the model calculations on the computer and the movement of the arm. The real-time classification of musical features is converted into electrical signals which are read by the robotic arm to control the movement patterns of the different effects. The arm converts the electrical signals it reads into coordinate system values to control the range and behaviour of the arm's movements (see Figure 3).

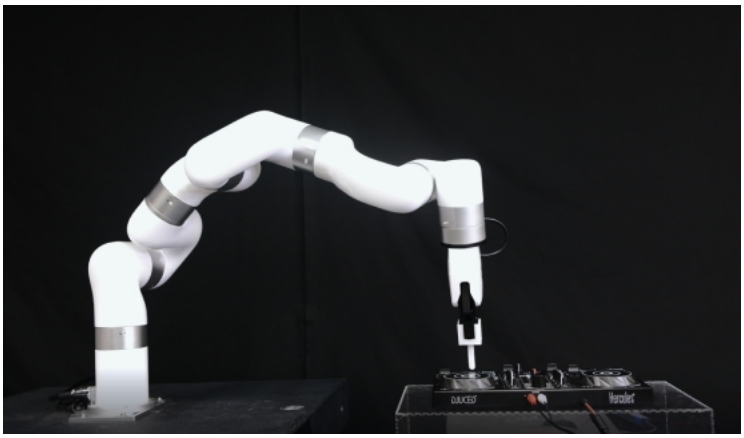

**Figure 3:** Operation of the robotic arm.

According to the previous steps, the users are allowed to operate one side of the DJ console while the robotic arm can operate on the other side to add some effects to enrich the remixing.

## USER EXPERIENCE EXPERIMENT

According to the complete system described above, a user experience experiment was conducted to verify that both novice and experienced DJs can use the system and robotic arm to perform mixing operations. Five users participated in the experiment, including three novices and two experienced DJs. They used the system in its entirety, including learning basic operations on a virtual DJ console in VR, watching others' works to express their preferences, and collaborating with the robotic arm to create music. In addition, three listeners independently evaluated the music created by the users and the mixed

songs created in collaboration with the robotic arm during the experiment to verify whether the music performances created by human and machine collaboration are more attractive to users. After completing the entire experiment process, users completed a user experience feedback form to express their satisfaction with the system and their understanding of collaborative DJ music production.

Novice DJs were the first to participate in the experiment. They first learned some basic mixing and effects operations on a virtual DJ console in VR. After understanding these operations, they selected one to two songs to create music on a physical DJ console. Then, the robotic arm was added to collaborate with the user in mixing operations. During this process, the audience watched the user's mixing performance and evaluated it based on five dimensions: sense of rhythm, transition, creativity, stage presence, and interaction. Based on these five dimensions, we defined five levels: poor, fair, good, outstanding, and excellent. The complete audience evaluation form (see Table 1). Listeners could check the corresponding evaluation criteria to complete the evaluation of the user's performance in the experiment. After completing the user experience test, users completed a Google form to provide feedback on the system.

**Table 1.** Audiencee evaluation on musical performance.

| Criteria | Poor | Fair | Good | Excellent | Outstanding |
|---|---|---|---|---|---|
| Sense of rhythm | | | | | |
| Transition | | | | | |
| Creativity | | | | | |
| Stage presence | | | | | |
| Interaction | | | | | |

Experienced DJs also followed the same process. As they already understood the functions and effects of pads on the DJ console, they could create DJ music in the VR environment and appreciate the works of others in the DJ community in VR while expressing their preferences. Then, they repeated the same operation as the novice DJs, but they could choose to create two to three songs, and the robotic arm's collaboration was based on the different creative inspiration generated by the user's preferences for music in the VR environment. Finally, the audience evaluated the music created by the users independently, and the users provided user feedback based on a personalized music collaboration experience. Both novice and experienced DJ feedback on the system will be presented in the next section.

## RESULTS

The user experience tests verified the effectiveness of the system, which reflected in the evaluation from audience. In comparison to the performance individually done by users, the performances collaborating with robotic arms were more successful as a whole shown in the evaluation table. The result

(see Figure 4 and Figure 5) shows that the robotic arm can assist users in capturing the rhythm points more easily, smoothly transitioning between songs, making the music sound more natural and seamless, and also inspiring the desire to dance in the audience. This was particularly evident in the criteria named sense of rhythm, and transitions, which revealed that this system could provide support in the field of music. In other words, the model defined in the system was functional. However, the impact of the robotic arm on the testers' creativity was not obvious since, in very limited time, the new music genres and modes that the system has brought forth were hard for users to comprehend. For factors that are not strongly associated with music which is stage presence and interaction, the impact of the robotic arm on testers varied from person to person. In terms of collaborating with users to create music, the robotic arm was more helpful to beginners than to experienced DJs.

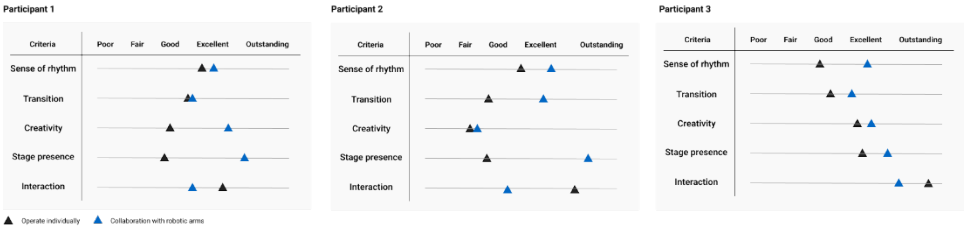


**Figure 4:** The audience evaluation on novice DJs performance.

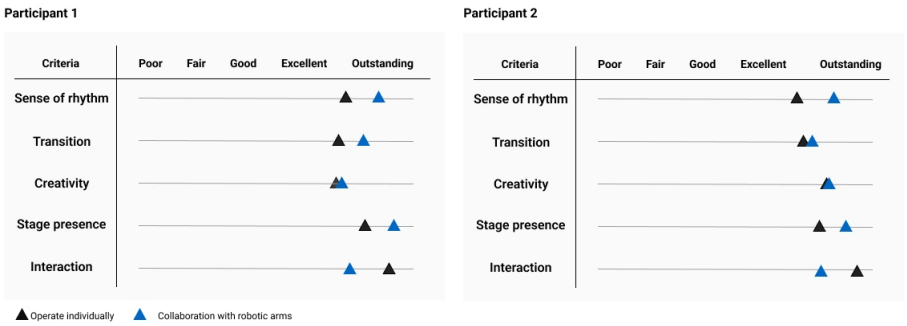


**Figure 5:** The audience evaluation on experienced DJs performance.

The audience assessment had already validated that this system was practical in musical creation. The users' complete experience with the system further validates its effectiveness in music creation. The robotic arm can assist users in analyzing the melody and rhythm of the music, as well as in the transition between songs, without colliding with the user's limbs during the process. Every user had a favourable impression of the whole system and described the experience as smooth and hassle-free. For both novice DJs and experienced DJs, they expressed satisfaction with this collaboration system in 7-point matrix. The average satisfaction was 5.4 out of 7, which indicating that the system was able to capture drumbeats and create synchronized and powerful DJ sets to assist users according to the feedback.

For novice DJs, due to the lack of expertise in music, the operations of robotic arms in dealing with musical features were fresh concepts to them, but they still had a distance to travel before fully comprehending it (see Figure 6). State differently, there was still a long way to go before they could independently apply these techniques in the field. As for experienced DJs, all of them considered the rhythmic patterns performed by the robot arm to be rhythmic and joyful. For the unique section of personalized remixing styles, the connection between the robotic arm's performance and the preferences expressed in VR is still unclear to current users. Experienced DJs tend to rate the robotic arm's music proficiency higher than novice DJs because of their extensive knowledge. However, the inspiration they received from the robotic arm's performance was not as direct as that of beginners since they already have their own preferences for suitable mixing points. Nevertheless, the robotic arm's mixing operation is more novel for experienced DJs, which can disrupt their thinking patterns to some extent. The vast majority of testers found that the robotic arm's manipulation of the music rhythm had a discernible rhythmic pattern.

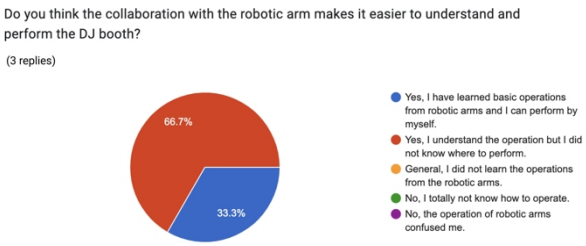


**Figure 6:** Questionnaire results on novice DJs on collaborative experience.

To sum up, for beginners, the robotic arm stimulated their interest and made them willing to continue learning DJ in the future. Experienced DJ users suggested that using VR to collaborate with them in trying new styles of music creation is also a good option. Despite the limited sample size in our experiment, machine learning's recognition of music beats is not precise enough due to the influence of melody and vocals, which can result in the robotic arm's identification of music characteristics and addition of mixing effects that may be unclear to users. Additionally, there is a delay in the ML model's recognition of music and its translation to the robotic arm's motion, which affects the performance to some extent. Furthermore, we also understand that a DJ's performance involves music, bodily movements, and interaction with the audience (Chang et al., 2019). The vision effects are also made of the whole performance (Vuoskoski et al., 2016). Therefore, the robotic arm could potentially incorporate new features to fully complement a DJ's performance. They are not just music partners, but their flexible joints and range of movement provide possibilities for dance and performance atmosphere. Overall, our system provides an opportunity for human and machine collaboration to create mixed songs and perform for the audience.

## ACKNOWLEDGMENT

This project was funded by the Intelligent Systems Design Programme, School of Design, the Hong Kong Polytechnic University. The work was also substantially supported by the Projects of Strategic Importance of The Hong Kong Polytechnic University (Project ID: P0036851). The authors would like to acknowledge the technical and ideation supports from Zhengtao Ma, Yaqi Zhang and Cong Fang at the Intelligent Systems Design Programme, School of Design, the Hong Kong Polytechnic University.

## REFERENCES

Beamish, Timothy, Karon Maclean, and Sidney Fels (2004). "Manipulating music: multimodal interaction for DJs". In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 327–334.

Callander, Mike (2022). "Thinking About Syncing: Examining the impact of 21st century DJ technology on the production and performance of Electronic Dance Music." In: Chroma: Journal of the Australasian Computer Music Association 38.1.

Chang, Andrew et al. (2019). "Body sway reflects joint emotional expression in music ensemble performance". In: Scientific Reports 9.1, p. 205.

Cliff, Dave (2000). "Hang the DJ: Automatic sequencing and seamless mixing of dance-music tracks". In: Hp Laboratories Technical Report Hpl 104.

Dis, Eva AM van et al. (2023). "ChatGPT: five priorities for research". In: Nature 614.7947, pp. 224–226.

Dwivedi, Yogesh K et al. (2023). ""So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy". In: International Journal of Information Management 71, p. 102642. 1

Huang, Hao-Wei et al. (2022). "AI DJ System for Electronic Dance Music". In: 2022 International Symposium on Electronics and Smart Devices (ISESD). IEEE, pp. 1–6.

Kumar, Yogesh, Komalpreet Kaur, and Gurpreet Singh (2020). "Machine learning aspects and its applications towards different research areas". In: 2020 International conference on computation, automation and knowledge management (ICCAKM). IEEE, pp. 150–156.

Montano, Ed (2010). "'How do you know he's not playing Pac-Man while he's supposed to be DJing?': technology, formats and the digital future of DJ culture". In: Popular Music 29.3, pp. 397–416.

Schäfer, Thomas and Peter Sedlmeier (2010). "What makes us like music? Determinants of music preference." In: Psychology of Aesthetics, Creativity, and the Arts 4.4, p. 223.

Schwarz, Diemo and Dominique Fourer (2021). "Methods and datasets for DJmix reverse engineering". In: Perception, Representations, Image, Sound, Music: 14th International Symposium, CMMR 2019, Marseille, France, October 14–18, 2019, Revised Selected Papers 14. Springer, pp. 31–47.

Tan, Xu and Xiaobing Li (2021). "A tutorial on AI music composition". In: Proceedings of the 29th ACM international conference on multimedia, pp. 5678–5680.

Tigre Moura, Francisco and Charlotte Maw (2021). "Artificial intelligence became Beethoven: how do listeners and music professionals perceive artificially composed music?" In: Journal of Consumer Marketing 38.2, pp. 137–146.
Vuoskoski, Jonna K et al. (2016). "Interaction of sight and sound in the perception and experience of musical performance". In: Music Perception: An Interdisciplinary Journal 33.4, pp. 457–471.
Wu, Xingjiao et al. (2022). "A survey of human-in-the-loop for machine learn.