Augmented Personification of Intelligent Music Tools for Creativity and Collaboration

Torin Hopkins
torin.hopkins@colorado.edu
ATLAS Institute,
University of Colorado Boulder

Rishi Vanukuru
rishi.vanukuru@colorado.edu
ATLAS Institute,
University of Colorado Boulder

Suibi Che-Chuan Weng
suibi.weng@colorado.edu
ATLAS Institute,
University of Colorado Boulder

Amy Banic*
amy.banic@colorado.edu
ATLAS Institute, University of
Colorado Boulder

Ellen Yi-Luen Do
ellen.do@colorado.edu
ATLAS Institute,
University of Colorado Boulder

Figure 1: Current intelligent music tools (A,D), our vision for personified interfaces (B,E), glimpses of our AR prototypes (C,F).

ABSTRACT

Intelligent Music Tools often aim to replicate human creativity, but the experience of interacting with them is limited by the interfaces of today. Such tools can be considered as collaborators in the creative process, and extending this metaphor to other modalities beyond audio, we propose the idea of personifying these tools as humanoid avatars in Augmented Reality to enhance the process of composing and performing music. In this paper, we present two types of personified interfaces based on generative music and live looping. Through this workshop, we seek to gain feedback on our ideas and prototypes, and potentially spark a discussion about future multi-sensory interfaces for collaborating with intelligent music tools.

KEYWORDS

Augmented Reality, Generative Music, Looping, Collaboration

ACM Reference Format:
Torin Hopkins, Rishi Vanukuru, Suibi Che-Chuan Weng, Amy Banic, and Ellen Yi-Luen Do. 2022. Augmented Personification of Intelligent Music Tools for Creativity and Collaboration. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29–May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

*Also with Interactive Realities Lab, University of Wyoming

CHI '22, April 29–May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 MOTIVATION

Musical Instruments—both physical and software-based—have traditionally functioned primarily as tools or conduits for the expression of human creativity. When considering musical tools that might be intelligent by themselves, the dominant paradigm has been that of "generative" entities powered by machine learning algorithms and models. A large majority of ML-powered intelligent

musical systems try to recreate the behavior and style of specific human musicians. In contrast to generation, the idea of "replicating" musical knowledge or behavior has been employed by musicians for many years in the form of live looping. Looping involves the recording and playback of musical phrases that are synchronized with each other. Musicians make a conscious attempt to express their creativity and store it in a tool that functions as a version of themselves, which then performs a specific role in the overall piece of art. One of the greatest amplifiers of musical creativity is collaboration, and both these ideas—generative music driven by machine learning, and live-looped music driven by human intelligence—can be seen as attempts at creating intelligent "collaborators" for musicians to work with. In the real world, musical collaboration in the form of live improvisational performances and jam sessions is an inherently multi-sensory activity, with visuals, audio, and haptics all playing an essential role in fostering a sense of collective creation. The same cannot be said for live looping or generative music tools, which for the most part exist as two-dimensional windows on a screen (Figure 1D), or a series of foot-pedals on the floor (Figure 1A). Our research seeks to question this status quo by asking, what if musical tools that run on human (or human-like) intelligence could be personified using Augmented Reality, and thus provide more of an embodied experience? To that end, we have created a series of prototypes which demonstrate how virtual avatars can be used to enhance the experience of working with intelligent musical tools.

2 RELATED WORK

We draw from literature in the fields of robotic musicianship and software-driven musical generation—two methods of achieving intelligent musical interaction that have been developing in parallel. Weinberg's Shimon, an autonomous robotic musician, improvises jazz music on the marimba and sings [2]. Earlier versions of robotics have been shown to run on algorithms generated in the music software Max/MSP[1]. Max/MSP is still used today to manage various low-level musical tasks, but has more recently been combined with advanced computational methods, creating robust musical information retrieval methods and generative musical output. These models have limitations with regard to processing speed, causing many researchers to favor a note-level model of generative output [7]. In contrast, software-driven musical exchange has been explored by several research and AI development teams [3, 5–7]. AI Duet, an AI-driven musician, plays short melodies in response to user piano keyboard input [4]. This has also been accomplished with the neural network transformer GPT-2 and GPT-3 [1]. The simplicity of the applications and the minimal software requirements make this approach highly accessible. However, there is an inherent trade-off between these two approaches. Robotic musicians are able to replicate the feeling of "presence" imparted by real collaborators, but software-based systems are far more accessible and generative in nature. The use of Augmented Reality to impart personified characteristics to software tools is an interesting approach that might help bridge this gap.

[1]https://cycling74.com/products/max

3 AR PERSONIFICATION PROTOTYPES

To explore the idea of personified interfaces, we created two distinct prototypes for the tools of live-looping (Figure 1B), and generative music (Figure 1E). The augmented reality experience was developed in the Unity live development engine for the Nreal Light headset[2]. The intelligent personifications are represented by AR avatars that appear to be in the same space as the musician. At present, we chose the case of musical input via a hand-drum in order to focus on rhythmic patterns before dealing with melody. The visual behaviour of these avatars was primarily controlled by MIDI information (either as part of the loop or generated by an algorithm) that triggered certain prerecorded animations that were collected from a motion-tracked live drumming session.

3.1 Live Looping

Here, musicians use a traditional foot pedal to indicate the beginning and end of looped sections. However, in addition to the audio being played back, our system places a virtual clone of the musician in front of them in AR, that physically plays the recorded rhythm (Figure 1C). As more loops are added, more avatars are spawned, almost recreating the feeling of being in a drum circle.

3.2 Generative Music

In the generative scenario, the AI virtual musicians are similarly represented by humanoid avatars in AR, and are driven by a combination of MAX for Live, Python, and Magenta Neural Networks[3]. The pipeline begins with a human musician playing a MIDI-output-capable instrument connected to a local computer. We manage MIDI information and sound production using MAX for Live and Ableton Live respectively. MAX for Live is used to read files from the player and pass the gathered data to a neural network (provided by Google Magenta). The neural network generates musical information in response to the player and saves output to local files. Local files are then read by MAX for Live and output using Ableton Live, which then in turn animates the personified representation of the algorithm (Figure 1F). We are currently working on making this pipeline more efficient in order to produce near real-time musical and visual response to a player's input.

4 CONCLUSION

Having developed these personified prototypes, we plan to conduct user tests to study how musicians' usage of the underlying tool changes when the virtual musicians are present. We also hope to incorporate more instruments into our pipeline, such as electronic keyboards and guitars. This work was inspired in part by the lack of in-person jamming opportunities during the COVID-19 pandemic, and an interesting next step would be to create networked sessions involving real musicians collaborating with each other remotely while also using their tools as personified avatars in the shared AR space. We are excited to learn more about the creation and processes behind intelligent musical tools from other participants at the workshop, and hope to bring our expertise in AR/VR experience design to help begin a discussion on how musical interfaces can move into the spatial world around us via immersive computing.

[2]https://www.nreal.ai/light/
[3]https://magenta.tensorflow.org/

ACKNOWLEDGMENTS

This material is based upon work supported in part by a grant from Ericsson Technology Office Silicon Valley.

REFERENCES

[1] Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, and Ilya Sutskever. 2020. Jukebox: A generative model for music. *arXiv preprint arXiv:2005.00341* (2020).

[2] Guy Hoffman and Gil Weinberg. 2010. Shimon: an interactive improvisational robotic marimba player. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. 3097–3102.

[3] Cheng-Zhi Anna Huang, Curtis Hawthorne, Adam Roberts, Monica Dinculescu, James Wexler, Leon Hong, and Jacob Howcroft. 2019. The bach doodle: Approachable music composition with machine learning at scale. *arXiv preprint arXiv:1907.06637* (2019).

[4] Matthew Hutson. 2017. How Google is making music with artifcial intelligence. (2017).

[5] Bill Manaris and Andrew R Brown. 2014. *Making music with computers: Creative programming in Python*. CRC Press.

[6] Jon McCormack, Toby Gifford, Patrick Hutchings, Maria Teresa Llano Rodriguez, Matthew Yee-King, and Mark d'Inverno. 2019. In a silent way: Communication between ai and improvising musicians beyond sound. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–11.

[7] Gil Weinberg, Mason Bretan, Guy Hoffman, and Scott Driscoll. 2020. *Robotic musicianship: embodied artificial creativity and mechatronic musical expression*. Vol. 8. Springer Nature.

.