CHIME One Day Workshop 2023

The Open University, Milton Keynes
4 December 2023

Bringing Timbral Shapes to Interactive Music Systems

Andrea Bolzoni[1*], Alistair Willis[1], David Sharp[2], Adam Linson[1], and Robin Laney[1]

[1]School of Computing and Communications, The Open University, UK, andrea.bolzoni@open.ac.uk
[2]School of Engineering and Innovation, The Open University, UK

***Abstract*— Artificial Intelligence has been successfully used in systems that imitate music style. However, when the systems take into consideration timbre features, they rely on a predefined sound database. We consider interactive systems that help musicians explore their style and we build on neural-based sound synthesis to support timbral music-making.**

I. Introduction

Since the 1950s, substantial research has been done on building co-creative music systems. Some of them are designed to imitate the style of the user they interact with, and we consider the Continuator [1] as an archetype of such systems. In this system, style is modelled on the pitch, duration, and intensity of the notes. Reflexive Looper [2] expands on the Continuator, integrating aspects of timbre using audio descriptors [3] and concatenative sound synthesis [4] for interactive audio generation. Yet, notes still have an important role in modelling style. Systems such as MASOM [5] generate sounds with concatenative sound synthesis, and analyse and model an audio corpus uniquely using a collection of audio descriptors.

II. Proposal

Theories such as spectromorphology [6] can help us analyse music practices that make music with timbre. A fundamental aspect is the shape - morphology - of the basic units of a music corpus. To elicit a sonic example, let's think of a mouth harp piece: the shape of each unit is as important in defining a musician's style as their sequencing is. More precisely, the shape of one or more timbre dimensions is relevant in defining a musician's style.

For illustration purposes, let's consider an instrument made only of a noise generator and a band-pass filter. The timbre dimensions we will make music with are the central frequency and bandwidth of the spectrum. To model the style of music played with this instrument, we will look at the shapes of the central frequency and bandwidth. And, if we want to generate music that is consistent with the modelled style, we want to generate the shapes without being bound to a fixed database of sounds.

Interactive music systems designed to explore a musician's style lack the capability to generate audio streams with shapes in timbre dimensions. We propose a model (Figure 1) that will integrate the generation of morphologies in interactive music systems, to support timbral music-making. This generative model will make it possible to expand the database of sounds from which the concatenative sound synthesis draws.

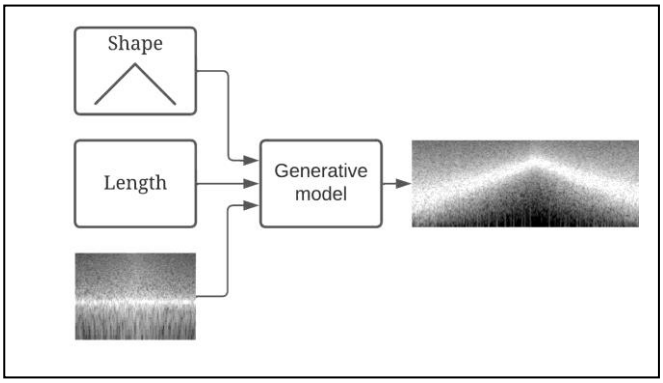

Figure 1. Generative model that can modify a timbre dimension and the duration of an audio input. The input is provided as a magnitude spectrogram. For illustration purposes, in this image we use filtered noise. The central frequency of the output takes the shape provided to the model.

We argue that the possibility of controlling shapes in timbre dimensions will make the output stream of an interactive system more coherent with the modelled style. Additionally, it will open up new ways to expand and explore music style. We are designing a study to evaluate to what extent the stream generated by our model is perceived as consistent with the style of the original audio corpus.

References

[1] F. Pachet, "The continuator: Musical interaction with style," *Journal of New Music Research*, vol. 32, pp. 333–341, 9 2003.
[2] F. Pachet, P. Roy, J. Moreira, and M. d'Inverno, "Reflexive loopers for solo musical improvisation," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013.
[3] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The timbre toolbox: Extracting audio descriptors from musical signals," The Journal of the Acoustical Society of America, vol. 130, pp. 2902–2916, 2011.
[4] D. Schwarz, "A system for data-driven concatenative sound synthesis," *Digital Audio Effects (DAFx)*, 2000, pp. 97–102.
[5] K. Tatar and P. Pasquier, "Masom: A musical agent architecture based on self-organizing maps, affective computing, and variable Markov models," *Proceedings of the 5th International Workshop on Musical Metacreation (MUME 2017)*, 2017.
[6] D. Smalley, "Spectromorphology: explaining sound-shapes," Organised Sound, vol. 2, pp. 107–126, 8 1997.

*Research funded by The Open University.