

PlayTrainPlay

MARK HANSLIP, University of York, UK

Additional Key Words and Phrases: improvisation, neural networks, interaction, granular synthesis

ACM Reference Format:

Mark Hanslip. 2024. PlayTrainPlay. 1, 1 (September 2024), 3 pages. <https://doi.org/10.1145/XXXXXXX.XXXXXXX>

1 PROJECT DESCRIPTION

PlayTrainPlay is a structured solo improvisation in two parts. The performance expands on my practice as an improvising saxophonist by combining it with granular synthesis, machine listening and neural networks. The performance structure is defined by the process of collecting data, training a neural network and interacting with the resulting model; in another sense, the structure describes a transition from a physical effect interface to a data-based one.

In the first section, I interact with a real-time granular effect, mapping the effect's parameters via a foot controller consisting of a tactile 2D surface and a switch. When pressed, the switch triggers MFCC analysis of the saxophone's timbre, logging the coefficients as input data and logging the current effect parameters as target data. Over the course of this section, an input dataset and a target dataset are thus built up. Towards the end of the section, the process of training the neural network - a shallow MLP-based autoencoder that learns mappings between the MFCC analysis and its corresponding granular synth parameters - is triggered via a second foot switch. Training is stopped by the performer when the loss no longer improves significantly, signalling the end of the first section.

In the second section, the performer no longer uses the foot controllers and instead relinquishes control of effects parameters to the neural network. The real-time effect from section 1 is replaced with an interactive loop in which segments of the player's input are sporadically recorded. Each segment's MFCC coefficients are extracted offline and passed to the trained neural network's weight space, transforming the MFCC data into effect parameters which inform the granular synth's response. The second section continues in this manner until the end of the piece.

This use of MFCC data for analysis of improvised saxophone timbre in order to shape an interactive system's generated responses has precedent in Bill Hsu's work with saxophonist John Butcher [1], whom I consider an influence. A differentiator between my proposed system and Hsu's (as well as other examples of machine learning for NIME such as Victor Shepardson's Living Looper [2], which incorporates pre-trained RAVE autoencoders [3]) is the performance of data collection and modelling as well as inference. The pairing of hardware controllers with granular synthesis software for transformation of an acoustic instrument takes inspiration from the work of Christian Fennesz [4], while diverging from it by allowing a neural network to control synthesis parameters: doing so affords a mid-point between user-defined parameters and surprising machine interventions. I also find the example of Jorrit Dijkstra's self-directed

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

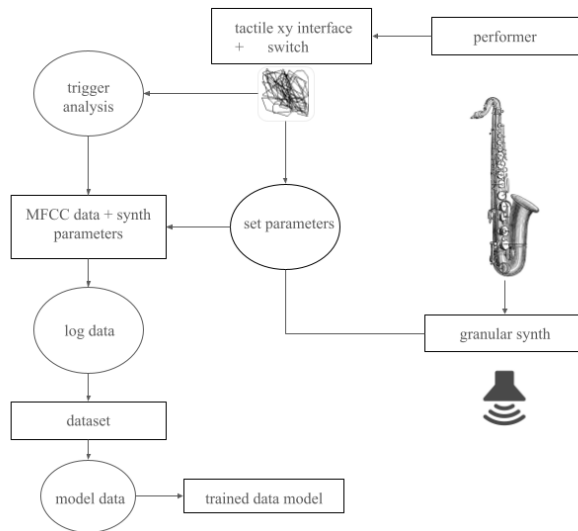
DOI: <https://doi.org/10.1145/XXXXXXX.XXXXXXX>

Music Proceedings of the International Conference on New Interfaces for Musical Expression

NIME'24, 4-6 September, 2024, Utrecht, The Netherlands

electronic augmentation of his saxophone improvisations highly motivating [5].

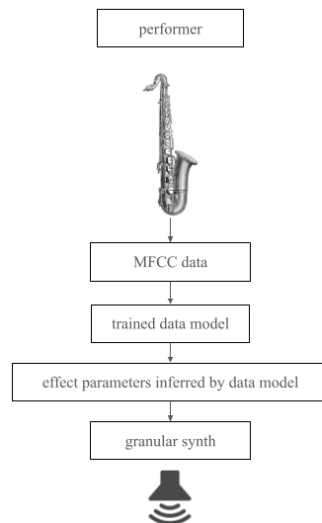
These processes are implemented in SuperCollider, using the FluCoMa toolkit [6] for machine listening and machine learning functionality and a small Python script for managing HID data. Section 1's model builds on Xenia Pestova's "External Hardware" model of interaction [7] by adding data logging and model training capabilities. Section 2's model is more streamlined and demonstrates the potential for data interfaces such as neural networks to simplify musicians' stage setups and workflows. The process flow of each section is outlined in the diagrams below.



Section 1

Improvisation through a real-time granular effect, controlling effect params via a tactile interface and periodically logging data via a footswitch.

Performer triggers data modelling via a second footswitch when ready.



Section 2

Performer interacts with the trained model.

The performer-controlled tactile interface has been replaced with a **data model**.

To enhance audience understanding and enjoyment, audio-reactive visuals showing inputs to and outputs of the computer program will be projected during the performance. The stage setup is intended to be as cable-free as is practical: both tactile controller and foot switches communicate with the laptop wirelessly, with only the audio and video outputs requiring cabling on stage, avoiding unnecessary entanglement.

REFERENCES

- [1] Hsu, Bill. 'Managing Gesture and Timbre for Analysis and Instrument Control in an Interactive Environment'. Proceedings of NIME 2006: International Conference on New Interfaces for Musical Expression, IRCAM - Centre Pompidou, Paris, France. 4 - 8 June, 2006.
- [2] Shepardson, Victor and Thor Magnusson. 'The Living Looper: Rethinking the Musical Loop as a Machine Action-Perception Loop'. Proceedings of NIME 2023: International Conference on New Interfaces for Musical Expression, Mexico City, Mexico. 31 May - 3 June, 2023.
- [3] Caillon, Antoine and Phillippe Esling. 'RAVE: A variational autoencoder for fast and high-quality neural audio synthesis', 2021. arXiv:2111.05011 [cs.LG].
- [4] Fennesz. *Endless Summer*. Vinyl, CD and Digital. Editions Mego, 2001. <https://fenneszemego.bandcamp.com/album/endless-summer>.
- [5] Dijkstra, Jorrit. *Never Odd or Even*. CD and Digital. Driff Records, 2015. <https://driffrecords.bandcamp.com/album/never-odd-or-even>.
- [6] Tremblay, Pierre Alexandre, Owen Green, Gerard Roma, Alexander Harker. 'From collections to corpora: Exploring sounds through fluid decomposition'. Proceedings of ICMC 2019: 45th International Computer Music Conference, New York, New York City, United States. 2019.
- [7] Pestova, Xenia. 'Models of Interaction in Works for Piano and Live Electronics'. PhD diss. Schulich School of Music, McGill University. pp21-24. 2008.