Real-time timbre mapping for synthesized percussive performance

Jordie Shier
C4DM, QMUL
j.m.shier@qmul.ac.uk

Supervised by:
Andrew McPherson
Charalampos Saitis

Industry Partner:
Andrew Robertson
Ableton

AI + MUSIC

Research Question

How can audio from acoustic percussion be mapped to controls for a drum synthesizer in real-time and in a musically expressive way?

Background

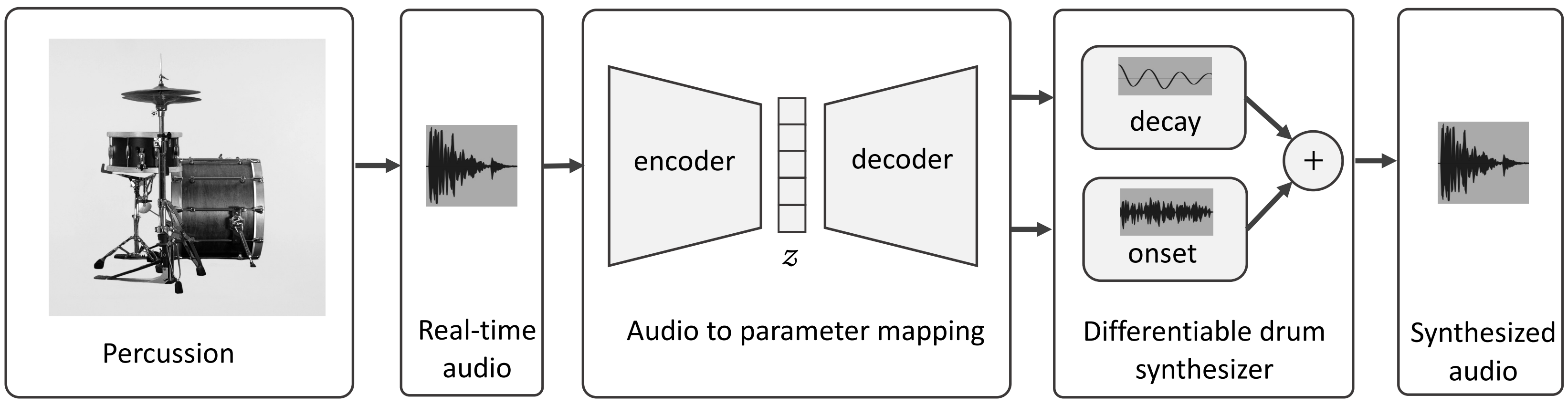
- Drum replacement using samples is common, but has limitations; augmenting drums using percussion synthesizers enables more expressive possibilities
- Synthesis of realistic percussive audio is active area of research [1]
- Timbre mapping applies the sonic qualities of one instrument to another while preserving aspects of the performance (i.e., timing, pitch, and loudness) [2]
- Research result will enable percussionists to create synthesized percussion tracks without having to learn a new instrument or purchase expensive sensors

Research Plan

1. **Differentiable Drum Synthesizer**: Build on techniques from [1] and implement differentiably [3], enabling backpropagation in ML frameworks
2. **Timbre and Parameter Mapping:** Explore generative networks for synthesizer parameter estimation [4], timbre mapping from acoustic percussion instruments, and timbral control interfaces [5]
3. **Real-time System:** Optimize implementation for real-time use and build on techniques for real-time timbre mapping of voice from [6]

References

[1] Kirby, Tim, and Mark Sandler. "The evolution of drum modes with strike intensity: Analysis and synthesis using the discrete cosine transform." The Journal of the Acoustical Society of America 150.1 (2021): 202-214
[2] Huang, Sicong, et al. "Timbretron: A wavenet (cyclegan (cqt (audio))) pipeline for musical timbre transfer." *arXiv preprint arXiv:1811.09620* (2018).
[3] Engel, Jesse, et al. "DDSP: Differentiable digital signal processing." *arXiv preprint arXiv:2001.04643* (2020).
[4] Esling, Philippe, et al. "Flow synthesizer: Universal audio synthesizer control with normalizing flows." *Applied Sciences* 10.1 (2019): 302.
[5] Javier Nistal, Stefan Lattner, and Gael Richard. Drumgan: Synthesis of drum sounds with timbral feature conditioning using generative adversarial networks. arXiv preprint arXiv:2008.12073, 2020.
[6] Stowell, Dan. *Making music through real-time voice timbre analysis: machine learning and timbral control.* Diss. Queen Mary University of London, 2010.
[7] Hayes, Ben, Charalampos Saitis, and György Fazekas. "Sinusoidal Frequency Estimation by Gradient Descent." *arXiv preprint arXiv:2210.14476* (2022).



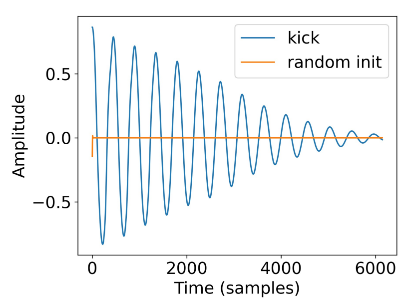
Proposed System
A network inspired by the DDSP autoencoder [3] is proposed as a method to map from real-time percussive audio to parameters of a differentiable drum synthesizer. The onset and decay portion of percussive sounds will be modelled separately within the synthesizer architecture.

Early Experiments

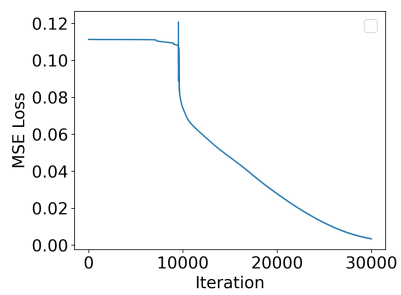
- Gradient descent was used to optimize the parameters of a simple differentiable kick drum to match the sustain of an electronic kick drum
- Kick modelled using a harmonic oscillator with exponentially decaying frequency based on [7]

$$y(t) = Ae^{-T_A t} * \cos(\omega e^{-T_\omega t} + \phi)$$

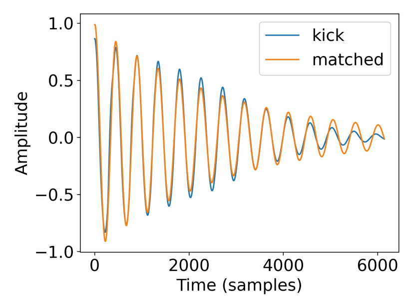
Target + Synth with Random Parameters



Mean Squared Error Waveform Loss



Target + Synth with Learned parameters



This work was supported by the UKRI and EPSRC under grant [grant number EP/S022694/1]

UK Research and Innovation

Ableton

Queen Mary University of London