

Audio Style Transfer

Laboratory:

Swiss Data Science Center

Type:

Semester Project

Description:

Style transfer between images consists of producing a picture that shares the content of an image (generally a photograph) and the style of another (typically a painting). This process has been studied extensively, and different approaches are known to produce appreciable results. On the other hand, the transfer of style (timbre) from an audio file to another has been substantially less explored, and the results achieved so far are not comparable in quality to those obtained with images.

In this project, we propose to replicate and compare some of the deep learning techniques experimented in the field of audio style transfer. To this end, we propose to exploit the available recordings of individual musical instruments as "content audio" and to take advantage of common effects (such as distortion, wah-wah, phaser, ...) to create "style audios". We will then develop an algorithm that transfers the effect from the "style audios" to the "content" ones.

The purpose of the project is to show that an algorithm can model a sound effect and apply it to an audio recording. If it will be possible to develop an algorithm with good performance, further efforts will be aimed at making it run in real time.

As this is an exploratory project, there is no guarantee that we will be able to develop a model with satisfactory performance. However, this does not affect the evaluation of the project.

Goals/Benefits:

- Development of an audio file transfer algorithm
- Working with (De)Convolutional Networks and Spectrograms
- Experience with audio processing

Prerequisites:

- Python (Advanced)
- Machine Learning (Intermediate)
- Experience with Pytorch is a plus

Deliverables:

- Well-documented code
- Written report and oral presentation

References:

- [1] Bhaumik Choksi, Alisha Sawant and Swati Mali. "Style Transfer for Audio using Convolutional Neural Networks." International Journal of Computer Applications 175 (2017)
- [2] Eric Grinstein, Ngoc Duong, Alexey Ozerov, Patrick Pérez. "Audio style transfer." ICASSP 2018 - 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
- [3] Prateek Verma, Julius O. Smith. "Neural Style Transfer for Audio Spectrograms." 31st Conference on Neural Information Processing Systems (NIPS 2017)

Contact:

giulio.romanelli@epfl.ch