CS229 Project Proposal

Sam Xu, Joyce Xu, Eric Tang

TOTAL POINTS

5 / 5

QUESTION 1

1 Project Proposal 5 / 5

+ **5 pts** All looks good
+ **0 pts** Did not submit
+ **4 pts** Project not properly scoped
+ **4 pts** Not enough detail
+ **3 pts** Insufficient work -- please come see project

TA

✓ **+ 5 pts** **Project does not make use of any technique taught in class. Full marks are being given but please consider using more CS229 techniques.**

Your mentor: Hojat

Please also use methods taught in this class.

Project Title: Neural Music Synthesis
Project Category: Audio & Music
Team Members: Sam Xu (samx), Eric Tang (etang21), Joyce Xu (jexu)

Motivation

Creativity in AI is an extremely difficult goal to accomplish, and music composition falls into the category of creative tasks. We hope to generate music using audio embeddings from deep neural networks, and by doing so, further our understanding of computational creativity.

Past work on this area has resulted in novel vector embeddings for music, which invites further work in analyzing the latent structure of these embeddings. We aim to visualize and explore the arithmetic relationships between these audio embeddings, using tools such as Principal Component Visualization and t-sne visualization.

Method

Previous results in music generation have only seen limited success, largely due to the difficulties in capturing long-term structures in music. Recently, however, the use of novel network architectures, such as LSTMs and variational autoencoders (VAE's), have opened new possibilities for music generation.[1] We hope to use variational autoencoders to generate musical structures at many different timescales.

We hope to use the NSynth dataset which contains over 300,000 annotated musical notes for over a thousand different instruments.[2] The dataset contains high-quality music files that will help us get started. In addition, we will gather a corpus of music that we seek to emulate, preprocess this corpus, and use it to train our model. Our tentative plan for a music corpus is a collection of popular electronic/synth music without lyrics.

Intended Experiments

Using the architectures mentioned above and the NSynth dataset of annotated musical notes, we aim to generate music for human evaluation. As a baseline, we will train a model using a vanilla Recurrent Neural Network to generate music.

Previous papers in music generation have used a small number of human evaluators to compare the performance of new models with previous models, as measured by evaluators' personal preference. We plan to recruit several human testers to compare the acoustic quality of our baseline with our new model.

[1] https://ai.google/research/pubs/pub46119
[2] https://magenta.tensorflow.org/datasets/nsynth