Parameter automation for granular synthesis

author:

Karol Bakunowski

supervisor:

Dr. Michael Zbyszyński

Submitted as part of the requirements

for the award of the Degree in

Music Computing

of

Goldsmiths University of London.

Department of Computing

Goldsmiths University of London

March 29, 2019

Abstract

My research is about stuff.

It begins with a study of some stuff, and then some other stuff and things.

There is a 300-word limit on your abstract.

A good abstract explains in one line why the paper is important. It then goes on to give a summary of your major results, preferably couched in numbers with error limits. The final sentences explain the major implications of your work. A good abstract is concise, readable, and quantitative. Length should be 1-2 paragraphs, approx. 400 words. Absrtracts generally do not have citations. Information in title should not be repeated. Be explicit. Use numbers where appropriate. Answers to these questions should be found in the abstract: What did you do? Why did you do it? What question were you trying to answer? How did you do it? State methods. What did you learn? State major results. Why does it matter? Point out at least one significant implication.

Acknowledgements

Acknowledge all the things!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Aims and Objectives

The initial aim of the project was to implement a machine learning solution to the task of granular synthesizer programming based on sound matching. Consequently building a tool that would assist musicians in creating interesting sounds, provided an audio input to the system.

Over the course of the past couple months the project extended into more research-oriented directions, with a focus of finding the best possible audio descriptors for specific characteristics of sounds. In order to determine the most significant in judging similarity on the cognitive level.

The methods here concerned with describing audio differ from other approaches in literature (e.g Matthew) primarily in that sound is treated in a modular manner, where a description is made based on one characteristic for X number instances, and separate processes are run for each. Instead of trying to describe audio as a whole, the aim is to describe it as a combination of things (density, pitch, rhythm, amplitude)

Concretely my objectives are:

1. To build a tool that is helpful to artists in creating new sounds

2. To challenge the interaction between an artist and a preset as a starting point to synthesis

3. To achieve a response that is not only stimulating to the user but also differs from simply randomizing the parameter values

In measuring the project's success, the subjective sonic coherence and similarity of algorithm's outputs may be considered the best indicator (the human discriminator?), along with more quantitative analysis, such as comparing the audio descriptors on input and output sounds for example.

In (future chapters) I provide some critique on established techniques of assessment of these types of problems and ultimately conclude that some stuff in the best way of doing it.

1.1.1 Deliverables

to these ends hehhe this is what i made:

Some stuff about things.*[example-citation]* Some more things. Inline citation: Anne Author. Example Journal Paper Title. *Journal of Classic Examples*, 1(1):e1001745+, January 1970

Chapter 2

Literature Review

2.1 Problem background

- what is the problem

- some study that describes it directly, and is a general thing done on the main problem here
- programming synths based on sound matching

2.1.1 Audio descriptors

MFCC - description - most popular in speech synthesis

what other can i mention that are useful, and widely used, but not as popular? mention the ones that i think will be useful in this project - paired with a granulator.

2.1.2 Granulation

difficult problem as the sound heavily depends on the audio sampled

to accommodate for that - some universal audio descriptor have to be used that describe rhythm, density, pitch etc - so that the input sound can be 'molded' into what will resemble the original input sound to the system.

concatenative synthesis - how it could be beneficial to make that - at least mention it and some studies about it. mention CataRT maybe...

2.1.3 Machine Learning

a neural network could me used to build a granular synthesis parameter space from a sound, but teaching it to do that, and predicting it's behaviour is a difficult task as well as a very experimental approach, that is quite difficult to controll and predict.

instead, focusing on replicating certain aspects of the input sound seems like a more straightforward and predictable approach to this problem.

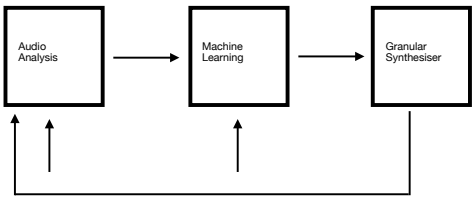
slicing the machine learning algorithm to predict different aspects of the sound separately will result in a more predictable outcome as well as one that's easier to control.

Chapter 3

Methods

The method of tackling this problem will be based on previous research done in this domain. There are three essential components for this project:

Figure 3.1: Basic flowchart



The box “Granular Synthesiser” in figure 1. represents a granular synthesizer. It will be build in C++, using openFrameworks and the Maximillian library.

“Machine Learning” in figure 1, stands for Machine Learning. This aspect will most likely be implemented in python, either using the “scikit.learn” library, or the “TensorFlow” library. Depending on which Machine Learning algorithms will be used. This will be determined during testing. The ones being considered at the moment include k-NN, Neural Networks, and LSTM(Long Short-Term Memory Neural Network)

The “Audio Analysis” box in figure 1. represents the aspect of the project responsible for extracting audio features from input sounds. This requires some testing in order to see which works best for this task. Feature extractors considered are: MFCC and DBM. Environment choices for now include the python “pyAudioAnalysis” for Python, and “Essentia” library for C++.

Training of whichever algorithm chosen will be happening on data generated by me. Vectors of synthesis engine parameter values, and audio analysis values, possibly as a CSV file will be created. The goal here is to automate a sampled walk through the parameter space, and extract audio features for each point. This was suggested to me by Dr. Rebecca Fiebrink, during her office hours, as part of feedback for another assignment. Possibly, a genetic algorithm could be used for this process

Once all the pieces are working, the key thing is to make them all work together as soon as possible, to create a working prototype, from which, by user testing and iterative design I can move forward and improve the project tackling issues one by one.

The minimum viable prototype has to include every aspect from Figure 1., excluding any substantial testing. Meaning, that one audio feature extractor, one machine learning algorithm, and a synthesizer that work together would be enough. Any testing of possibilities, whether audio feature extractors, or machine learning algorithms is outside of the scope of a prototype. As long as each piece of software can communicate and produce a desirable end result, the prototype will be considered successful.

Chapter 4

Results and Analysis

Chapter 5

Discussion

5.1 Summary of Findings

5.2 Evaluation

5.3 Future Work

Start with a few sentences that summarize the most important results. The discussion section should be a brief essay in itself, answering the following questions and caveats:

What are the major patterns in the observations? (Refer to spatial and temporal variations.) What are the relationships, trends and generalizations among the results? What are the exceptions to these patterns or generalizations? What are the likely causes (mechanisms) underlying these patterns resulting predictions? Is there agreement or disagreement with previous work? Interpret results in terms of background laid out in the introduction - what is the relationship of the present results to the original question? What is the implication of the present results for other unanswered questions in earth sciences, ecology, environmental policy, etc....? Multiple hypotheses: There are usually several possible explanations for results. Be careful to consider all of these rather than simply pushing your favorite one. If you can eliminate all but one, that is great, but often that is not possible with the data in hand. In that case you should give even treatment to the remaining possibilities, and try to indicate ways in which future work may lead to their discrimination. Avoid bandwagons: A special case of the above. Avoid jumping a currently fashionable point of view unless your

results really do strongly support them. What are the things we now know or understand that we didn't know or understand before the present work? Include the evidence or line of reasoning supporting each interpretation. What is the significance of the present results: why should we care?

This section should be rich in references to similar work and background needed to interpret results. However, interpretation/discussion section(s) are often too long and verbose. Is there material that does not contribute to one of the elements listed above? If so, this may be material that you will want to consider deleting or moving. Break up the section into logical segments by using subheads.

Chapter 6

Conclusion

What is the strongest and most important statement that you can make from your observations? If you met the reader at a meeting six months from now, what do you want them to remember about your paper? Refer back to problem posed, and describe the conclusions that you reached from carrying out this investigation, summarize new observations, new interpretations, and new insights that have resulted from the present work. Include the broader implications of your results. Do not repeat word for word the abstract, introduction or discussion.

Appendix A

An Appendix About Stuff

(stuff)

Appendix B

Another Appendix About Things

(things)

Appendix C

Colophon

This is a description of the tools you used to make your thesis. It helps people make future documents, reminds you, and looks good.

(example) This document was set in the Times Roman typeface using LaTeX and BibTeX, composed with a text editor.

Bibliography

[1] Anne Author. Example Journal Paper Title. *Journal of Classic Examples*, 1(1):e1001745+, January 1970.