|  |
| --- |
| Rochester Institute of Technology |
| Detecting Genre |
| What clues can programmers use to detect genre in audio programming |

|  |
| --- |
| Sean Levorse  April 3, 2016  12 E Main St. Denville, NJ 07834  (973)-555-1234  sml2565@rit.edu |

# Summary

Music is one of the largest forms of entertainment in the world. Music is often called the universal language. Generally, musical compositions can be classified into genres such as classical, rock, jazz, etc. However, these songs have to be classified manually by people and when songs start to blur the lines between two or more genres, classification can become messy. In this proposal, I outline the process I would like to research that might be able to quantify separating music into various genres in such a way that computers could be programmed to determine genre. I am a software engineering student at Rochester Institute of Technology with an app domain in Interactive Media. I have also taken a few classes that cover the physics of waves and sound as well as a music theory class taught by my advisor, Edward Schell. I have work experience at General Electric(GE) programming image processing and have studied signal processing enough to translate those concepts into audio processing. My education and work experiences combined with the guidance from Professor Schell covers the scope of this research.

# Introduction

One major field of audio signal processing is interpreting music. One issue with interpreting music is interpreting genre. Given just information about pitch and volume at a certain sampling rate, how can a programmer determine the genre of a song? This information is important for classifying music and would give the programmer more information for playlist generation. In order to observe how musical compositions compare to each other, I would use aspects from music theory such as tempo, key, and form. Another way to analyze genre is through the instruments. Detecting instruments through programming would require looking at the timbre[[1]](#footnote-1) of all of the instruments involved and using phasors to separate different additive frequencies[[2]](#footnote-2). The objective of this project is to tackle an issue in software and the methods needed to apply quantitative values to physical things we interpret in a qualitative manner. This project would fit into the interdisciplinary study category. It takes methods from music theory and the physics of waves and applies them to a programmatic context. I have taken courses in wave physics, interactive media, signal processing, and music theory. I also have been on co-op with GE doing image processing, a related signal processing field. This project would give me the opportunity to see how these fields interact.

# Methodology

Initially, I would need a few weeks to understand what genres are out there. Some of the obvious ones like rock and jazz wouldn’t be too hard to define, but the less common ones like tropical house would take more understanding. I would select a few genres, both broad and specific to analyze. From there, I would research what makes up those genres; is there a common form? Are songs generally major or minor key? What range do song tempos fall in? What kinds of instruments, if any, are used? From there, I could get to work on creating quantifiable values that would help determine if two songs are similar enough to be considered part of a genre. Once I have discrete methods, it would then be possible to create a program that takes two mp3 files and scores them based on if they are in the same genre. To do this, I would use C# in combination with the free and open sourced library CSCore. Then it would be possible to generate an Euler graph of songs to create an interconnected web of genres.

# Schedule

Week 1 – Research and define genre

Week 2 – Measure timbre, define discrete process for determining genre

Week 3 – Define methods for pulling out instruments from a song

Week 4 – Define concrete ranges for songs to be in the same genre

Week 5 – Define methods to sort songs by genre

Week 6 – Create program to populate database with genres, define genres

Weeks 7 and 8 – Create sample application (Smart Shuffle)

Weeks 9 and 10 – Buffer weeks for if one step takes longer than expected

# Materials

This project would require access to scholarly journals to complete the first step in the initial research. Depending on how well documented various instrument timbres are, there may be a need to calculate them. This would mean temporarily acquiring those instruments and a device to record the frequency of the waves emitted by the instrument. After this point, the only remaining factor would be legally acquiring mp3s for songs to test with.

# Budget

|  |  |
| --- | --- |
| Item | Cost |
| Scholarly Journals | $2000 |
| Instrument Rentals | $1000 |
| Audio Analyzer | $2000 |
| Mp3s | $200 |
| Salary | $4800 |
| Total | $10,000 |
|  |  |

# Advisor Biography

Edward Schell is a music theory professor at Rochester Institute of Technology. He has traveled the world studying music and the evolution of music throughout history. He knows how various genres have been born from older genres and how to distinguish one genre from another. He has written books on the evolution of electronic music and the history of music in the United States. This will come in handy as I predict electronic music may be one of the hardest genres to establish a timbre for. He can also clarify for me how genres are related to each other. His background in the study of music should complement my education in computer science, interactive media, and physics should cover all of the aspects of this project.

# Conclusion

Genre is one of the hardest aspects of music to quantify. As humans, we can distinguish songs that are similar, but it is difficult for a computer to distinguish between genres. This research would provide methods for programmers in the music industry to classify and sort musical compositions for their purposes. Marketing companies could use this information to see how different musical genres appeal to different demographics. The results of this research could be used by radio stations and online radio streams to produce effective playlists that don’t seem to jump around. The National institute for research should fund this project in order to improve the quality of entertainment across the entire music industry.

1. Timbre: Often called the “voice” of an instrument, timbre is what makes a C on a violin sound different than a C on a piano and a C on saxophone. They all have the same base frequency, but each instrument has a different set of resonant frequencies that accompany it. [↑](#footnote-ref-1)
2. Frequencies: The amount of rotations a sound wave goes through in a second. This is what determines the pitch of a certain sound. [↑](#footnote-ref-2)