

# Multimedia Information Retrieval System Report

Hugh Rawlinson

2008-1-13

## 1 Introduction

For this assignment, I decided to implement an audio based information retrieval system focused on short, percussive sounds that one might load into a sampler for use in the production of music.

## 2 Design decisions

### 2.1 Audio Analysis

I began by researching a variety of temporal and spectral metrics that are relevant to percussive sounds.[5] I chose to use Spectral Centroid, Spectral Flatness, and RMS as the features for this system, as well as the comparatively simpler sample time, as they're all directly related to obvious perceptual qualities of these types of sounds. The brightness, noisiness, loudness, and duration of short (less than 3000ms) percussive sounds are all qualities that are taken into major consideration when choosing appropriate percussive accompaniment in modern contemporary music.

There has been plenty of research into percussive information classification in the past[2][6][4], that seem to generally support the features I have chosen for this system.

### 2.2 Implementation tools

I chose to implement this system using a combination of R, for feature extraction, and Python, for human interaction and ranked lookup. I had originally considered implementing the system in C, for efficiency over large data sets, and entirely in Python for a cohesive codebase, but because of their lack of

easily accessible libraries for audio analysis (I had considered `essentia`[1] in C, but installation problems on my machine caused issues, as did a selection of libraries in Python), so I decided to use the tools I was most comfortable with for each of the individual tasks that the system has to perform.

### 3 Implementation

A short R script allowed me to extract the target audio features from a specified file on disk. From Python, I then wrote a command line utility to facilitate the process of creating a database of metrics of files within a directory, and to provide a simple cli for the user to lookup closest matching samples from their own dataset.

### 4 Evaluation

My evaluation of this project highlights some advantages of my approach, and reveals a set of improvements I would like to make. One advantage is that the set of queriable files is not limited - the user can run analysis on their own sets of data. In my testing, I have found that the system works quite well on my set of percussive sounds. The improvements that I would like to make include functionality to allow users to specify a set of database files based on a textual input listing file rather than relying on their filesystem to manage this database. At the very least, recursion into subfolders would be nice. I would like to do further research into the set of audio features for classification that are applicable to my target audio type, and implement further metrics to improve accuracy of the system. I intend to look into several ideas I have in hopes of making the system more cross platform. One particular feature that I would like to focus on is application to non-western percussive instruments. However, to accomplish this, I would need to evaluate the perceptual attributes of the sounds that are culturally significant to the home culture of each tradition[3], which is beyond the scope of this assignment. I also plan to create a graphical user interface for ease of use.

### 5 Summary

In summary, I have found this assignment to be incredibly interesting, and fully intend on continuing development on this project.

## References

- [1] D. Bogdanov, Nicolas Wack, Emilia Gómez, Sankalp Gulati, P. Herrera, O. Mayor, G. Roma, J. Salamon, J. Zapata, and Xavier Serra. Essentia: an audio analysis library for music information retrieval. In *International Society for Music Information Retrieval Conference (ISMIR'13)*, pages 493–498, Curitiba, Brazil, 04/11/2013 2013.
- [2] Ajay Kapur, Manj Benning, and George Tzanetakis. Query-by-beat-boxing: Music retrieval for the dj. In *Proceedings of the International Conference on Music Information Retrieval*, pages 170–177, 2004.
- [3] Thomas Lidy, Carlos N Silla Jr, Olmo Cornelis, Fabien Gouyon, Andreas Rauber, Celso AA Kaestner, and Alessandro L Koerich. On the suitability of state-of-the-art music information retrieval methods for analyzing, categorizing and accessing non-western and ethnic music collections. *Signal Processing*, 90(4):1032–1048, 2010.
- [4] Nobutaka Ono, Kenichi Miyamoto, Hirokazu Kameoka, and Shigeki Sagayama. A real-time equalizer of harmonic and percussive components in music signals. In *ISMIR*, pages 139–144, 2008.
- [5] Geoffroy Peeters. A large set of audio features for sound description (similarity and classification) in the CUIDADO project. Technical report, Icrum, 2004.
- [6] Adam R Tindale, Ajay Kapur, George Tzanetakis, and Ichiro Fujinaga. Retrieval of percussion gestures using timbre classification techniques. In *ISMIR*, 2004.