William Victor Simmons

St. Catharine's College

wvs22

Part II Individual Project Progress Report

Techniques for Separation of Harmonic Sound Sources

28 January 2017

Project Originator: William Victor Simmons

Project Supervisor: Dr D.J. Greaves

Director of Studies: Dr S.N. Taraskin

Overseers: Dr T.G. Griffin and Dr P. Lio

General Accomplishments

Two algorithms for uninformed, harmonic sound separation have been considered and successfully implemented. This has included producing working versions of the required numerical transforms (most significantly the Short-Time Fourier Transform and its inverse and Non-negative Matrix Factorisation by multiplicative-step gradient descent), Lloyd's algorithm for k-means clustering, file handling tools and bringing these all together with the actual separation algorithms.

In their current state, these solutions have shown that they can separate a mixture of two sounds from different musical instruments with reasonable likeness to the original sounds in waveform and envelope, provided that appropriate parameter values are selected for noise threshold, window size and hop size for taking the spectrogram and preference weightings for grouping sinusoids (how much the considerations of amplitude envelopes, frequency envelopes, harmony and stereo position affect the clustering).

Considerations of the Planned Schedule

At the time of writing this progress report, I have completed the objectives and produced the deliverables for all work segments due. High workloads from other aspects of the course towards the end of Michaelmas term led to the investigation and prototyping of the solution using Sinusoidal Modelling techniques was not completed until about four weeks after the initially proposed deadline. Consequently the investigation and prototyping of the NMF solution and rewriting of the first solution were not completed until two weeks after anticipated. The full implementation and testing of the second solution was, however,

completed on time. With workloads expected to be somewhat steadier during Lent term, I anticipate that setbacks like those experienced before should not be faced or will be less severe for the upcoming work segments.

I anticipate being in the position where I can consider developing and implementing a third solution as an extension later this term. I am also looking to continue further investigations on the current solutions as well. I am considering using soft-clustering for the sinusoids/spectrogram matrix factors to resemble how humans "hear out" sounds from a mixture and possibly improve the reconstruction quality. I would also like to experiment with other heuristics and methods for finding associated sinusoids/factors.