DrumBooth

A learning tool for drummers

Sean Breen R00070693 DCOM4
Wednesday 9th December 2015

Introduction

For my Final Year Project, I wanted to create an application that drummers could use to help with their learning by enhancing or suppressing the percussion in audio recordings.

Being a self-taught drummer myself for over 10 years, I've noticed that there is a lack of music tuition software where the focus is on the drums.

I want to provide drummers of all levels with the ability to help themselves improve using some standard features like a metronome and rudiments, but also help them learn to listen for the drums in music.

The ability to enhance the drums in an audio recording is very desirable. So, to do this, a method to separate the percussive and harmonic components from an audio signal was explored.

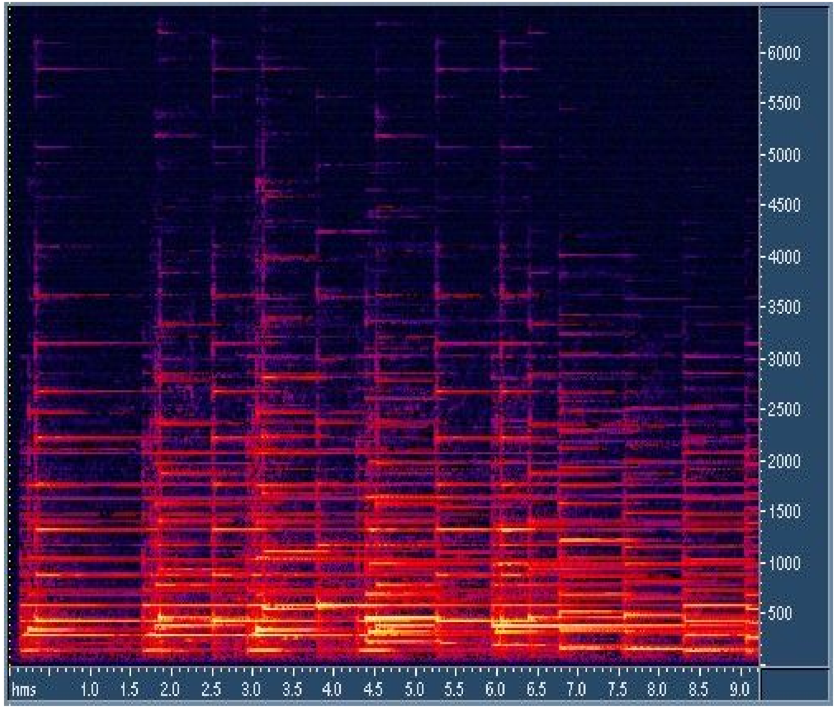
The algorithm I use is based on the method in Derry Fitzgerald's paper 'Harmonic/Percussive Separation using Median Filtering' from the International Conference on Digital Audio Effects (DAFx) 2010.

The Spectrogram

The spectrogram is a visual representation of an audio signal that describes the signal in terms of its energy of each of its frequency components and how they vary over time. In a computer, this is represented by a list of complex, floating-point numbers.

A digital signal's spectrogram is obtained by way of a Discrete Fourier Transform (DFT), which decomposes the signal into its individual frequency components. I use the Short Term Fourier Transform(STFT) algorithm to compute the DFT.

An inverse STFT is used to transform a spectrogram back to audio data.



Harmony and Percussion in Music

Music is composed of many different sounds. The components of music can loosely be described as being either **harmonic** or **percussive** in nature.

Harmonic elements would be notes played by a pitched instrument like a guitar or violin. These appear as horizontal lines on a spectrogram, that is, a stable frequency across time.

Percussive elements would be sounds created by hitting an instrument, such as the drums. These appear as vertical lines on a spectrogram. Vertical lines signify noisy sounds with no stable pitch.

However, these are not mutually exclusive. Many pitched instruments have percussive elements to their sounds and many percussive instruments have harmonic qualities. There are also several pitched percussion instruments, like the xylophone which create a harmonic sound by way of a percussive motion, i.e hitting the instrument with a mallet.

DrumBooth - Features

Application Features

- Tool to enhance or suppress percussion or harmony in a stereo or mono audio track.
- Programmable metronome.
- Browsable library of rudiments and examples.

Separation of Harmonic and Percussive components

The core of the application is the ability to enhance or suppress percussion in an audio recording.

The practical application is that somebody who wants to learn the drums to a song can load up the audio file and listen to it with the drums enhanced (to learn) and with the drums suppressed (to practice).

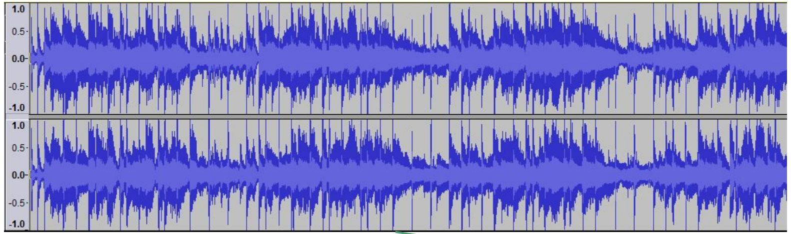
Aside from separating the harmony and percussion for learning, the files which are output by the application can be used to make remixes or backing tracks with either the percussive or harmonic elements enhanced.

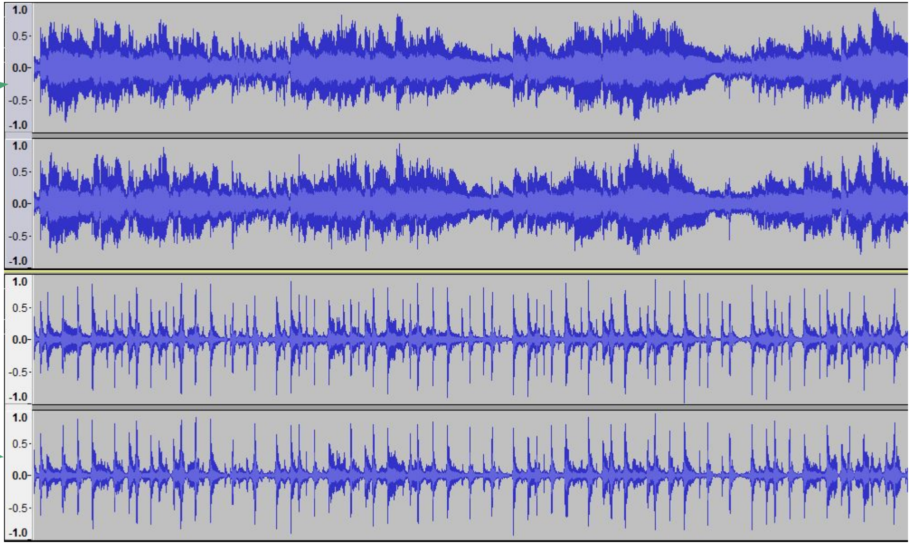
The filter I implemented uses a median filter technique across the frequency and time axes to get the harmony-enhanced and percussion-enhanced spectrograms, respectively.

Separation of Harmonic and Percussive components

Extracted Harmonics

Original audio file



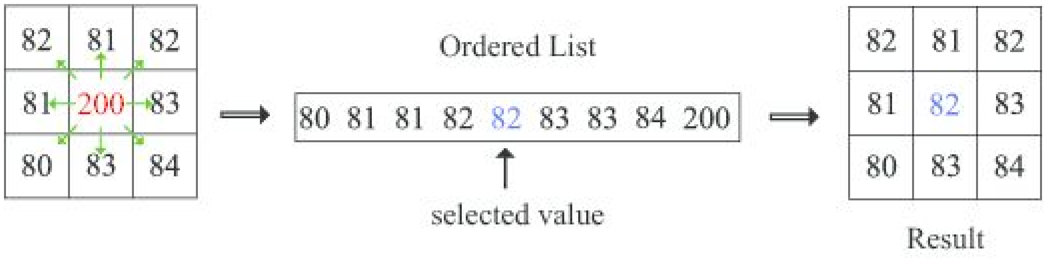


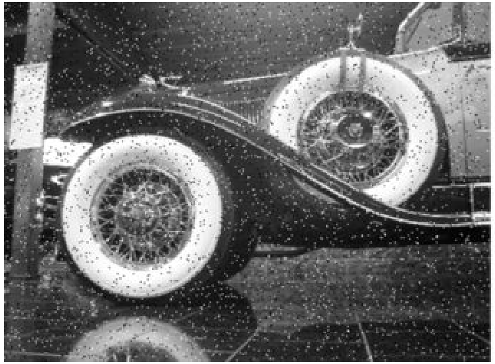
Extracted Percussion

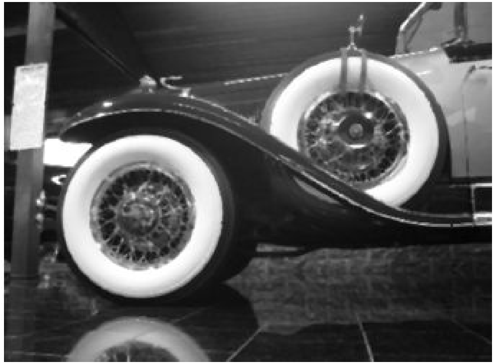
Median Filter

Median filtering is a signal processing technique usually used for image processing to remove noise from an image.

It works by assigning each pixel the median value of a number of neighbouring pixels. The neighbouring pixels are sorted by value, and the pixel is assigned the middle value. This is repeated for every pixel in the image.







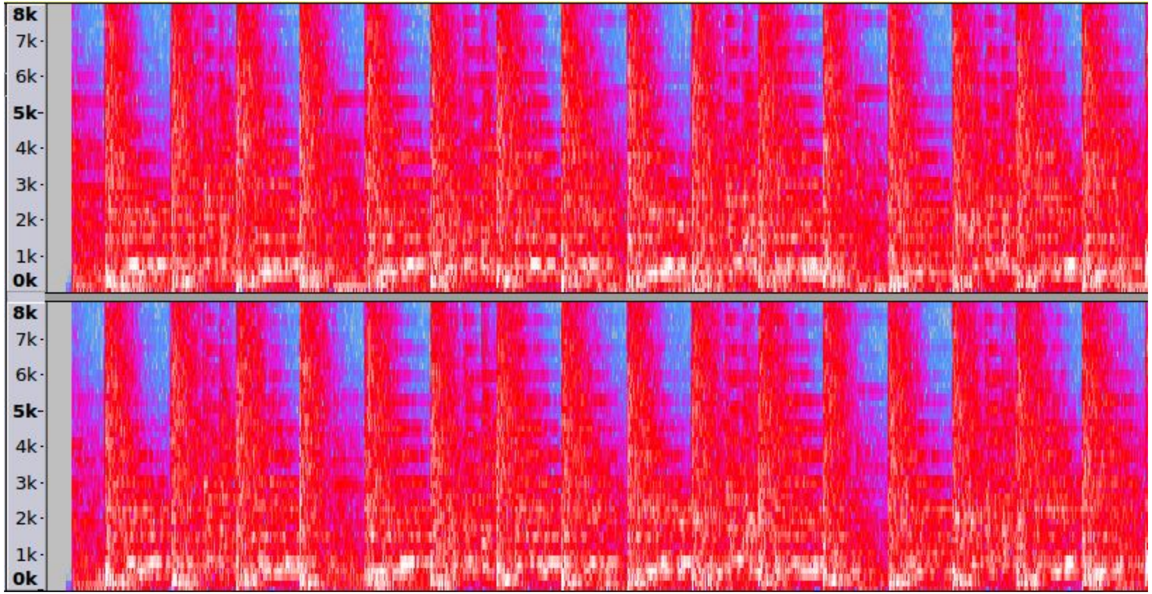
Example of a median filter that uses 8 neighbouring pixels.
http://blog.kleinproject.org/?p=588

Median Filtering on a Spectrogram

We can use the aforementioned technique on an audio signal's spectrogram to remove certain 'noise' from the signal.

The spectrogram is decomposed into time frames (columns) and frequency bins (rows).

We perform two loops, one on the time frames, and one on the frequency bins, and in each loop we perform the median filter on each sample, using a filter size of 17.



Excerpt from the beginning of Bruce Springsteen's 'Born in the U.S.A'

...continued

In each loop, we use the median filter to remove any ‘transients’ or spikes from the signal.

For each time frame, any tone with a stable frequency (horizontal line) are seen as a spike and will be removed with the median filter.

The same is true for each frequency bin, where any vertical line (percussion) are seen as a spike and will be removed.

After performing the two filter operations, we will be left with two spectrograms. The first spectrogram contains the percussion-enhanced audio, with the harmonic elements suppressed. The second spectrogram contains the harmonic-enhanced audio with the percussion suppressed.

These spectrograms are called ‘masks’ and will be applied to the original signal to separate the components.

...continued

This method has many advantages:

- Simple algorithm, easy to implement.
- It is very fast.
- Suitable for real-time audio processing.
- Not heavy on the CPU.

However, this technique does not provide perfect separation of drums from the rest of the instruments. The reason for this is that many instruments have percussive qualities, such as the sound of the pick hitting the strings of an acoustic guitar or the sound of rapidly changing vocal patterns, which will come out in the percussive spectrogram.

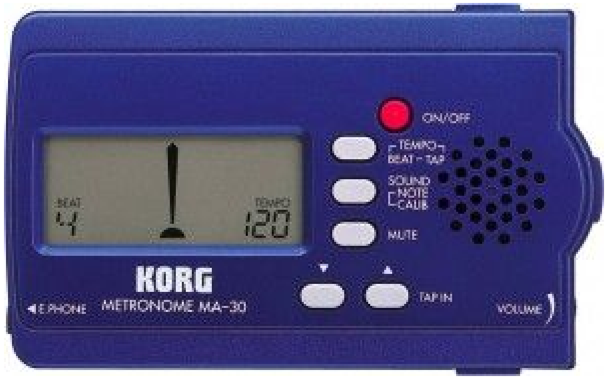
Additional filtering could be explored to achieve better results, however that would require more processing which would take more processing time and resources.

Metronome

DrumBooth has a metronome which can be enabled at any time. The metronome is used to provide a pulse to which a musician can practice their rhythm and timekeeping.

The metronome BPM and time signature can be programmed manually by the user.

In the future I intend to implement automatic beat detection and tempo tracking for the metronome.

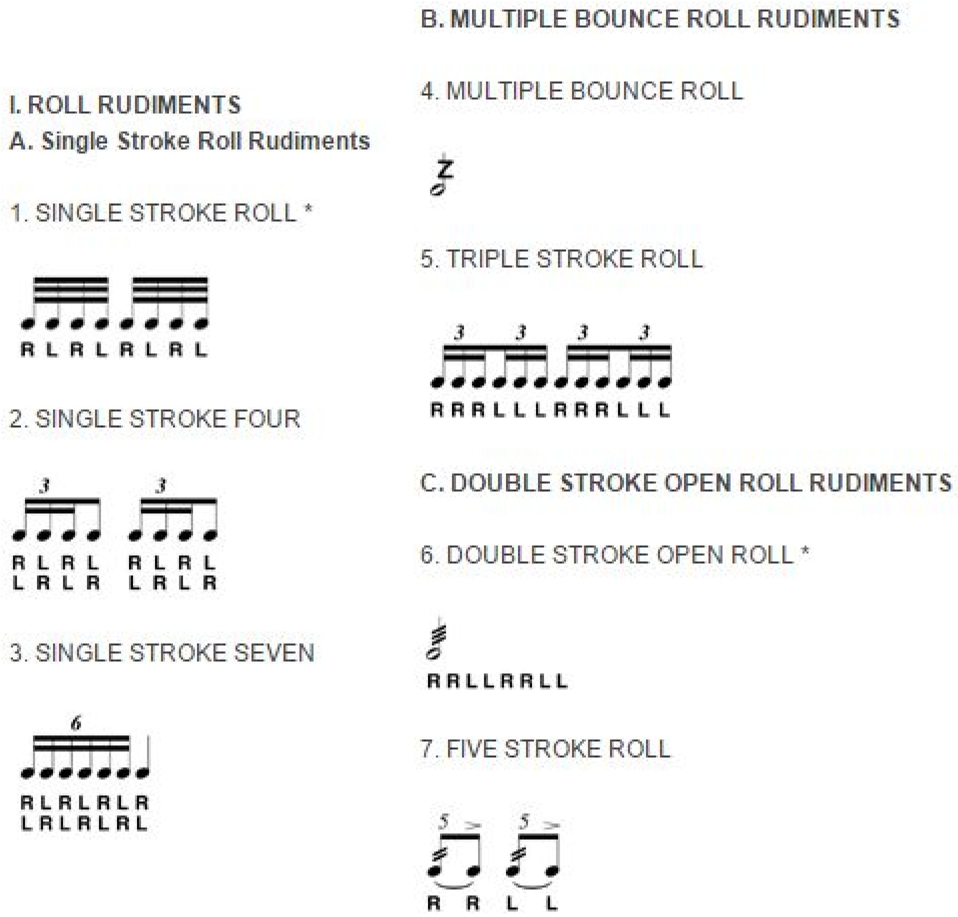


Rudiment Library

DrumBooth provides a browsable list of the 40 drum rudiments from the Percussive Arts Society.

The rudiment list includes the 26 American rudiments along with some contemporary, drum corps and orchestral rudiments.

These should be practiced with the metronome at various speeds and with both hands leading, i.e starting with both left and right hands.



Using DrumBooth for Learning

To use DrumBooth to learn a track, the user should pick an audio file which they want to learn the drums to and load it into the program.

The audio can be played like a regular audio player, but it has the added option to enhance or suppress the percussion in real time using a knob control where the centre point is the original track. Rotating the knob in either direction enhances either the harmonic or percussive while suppressing the other.

To practice effectively, the user should:

1. Move the knob control to suppress the harmonic elements and enhance the percussion.
2. Listen to the percussion-enhanced version of the track to hear the drums more clearly.
3. Once you feel you are familiar with the drum patterns, move the knob control to suppress the percussion.
4. Play along with the harmonic-enhanced version of the track.

Alternate Uses

The tracks which are output from the program can be saved and can be used to:

- Make remixes with the percussion or harmony enhanced.
- Make backing tracks to play along to.
- Make samples which can be used in other music.

This type of processing is also useful as a pre-processing stage for other audio applications, as in:

- ‘Improving Time-Scale Modification of Music Signals Using Harmonic-Percussive Separation’ Jonathan Driedger, Meinard Müller, Member, IEEE, and Sebastian Ewert
- ‘Signal Enhancement for Gunshot DOA Estimation with Median Filters‘ A. M. C. R. Borzino , J. A. Apolinario Jr. Military Institute of Engineering (IME), M. L. R. de Campos, L. W. P. Biscainho Federal University of Rio de Janeiro (UFRJ)

Prototype - Demonstration

Prototype

As a prototype, I decided to write a command-line application in C++ to show the effectiveness of this particular separation technique for percussive/harmonic separation.

Source code and an executable for this project prototype is available at

http://github.com/mangledjambon/drumbooth.

1. It takes a filename as an input parameter.
2. Performs a Short-Term Fourier Transform to turn it into a spectrogram.
3. Runs filter operation on every time frame to get percussive spectrogram.
4. Runs filter operation on every frequency bin to get harmonic spectrogram.
5. Applies the masks and re-synthesises the spectrograms back to audio data.
6. Writes the two files to the disk as stereo .wav files with a sample rate of 44.1 kHz.

Thanks for listening!