Onset Detection in Music

Statistical Methods IV End-Semestral Project

Tuhin Saha (BS2127)
Purushottam Saha (BS2119)

Instructor:	*Dr. Arnab Chakraborty*
Department:	ASU, ISI, Kolkata
Date:	January 13, 2024

Contents

1 Introduction

The field of automatic beat tracking in music excerpts is currently a thriving area of research. The objective is to develop a computational algorithm that can accurately identify the moments of beats in a musical excerpt in real-time, aiming to closely replicate the experience of human foot tapping. However, before we do analysis of the beats, we must first identify the onsets of musical events.

The automatic identification of events in audio signals opens up new possibilities in various music applications such as delivering content, compressing files, indexing, and retrieving music.

2 Methodology

There have been many different approaches for onset detection, e.g. taking ”novelty” curves based on the spectrogram, complex ”novelty” curves and many more. We, in particular, find an energy-based ”detection curve” or the ”novelty” curve, and finally detect peaks from this curve to get the required timestamps of the onsets.

Onset detection is the task of determining the starting times of notes or other musical events as they occur in a music recording. When such things happen, there is an energy jump in the signal, and by that, we can detect the note onsets. Here we try to find a local energy function and take the discrete differentiation to get the energy jumps. As we only care about the increase in energy, i.e. the note onsets, we use the half-wave rectification, by omitting the negative values in the energy difference/ energy jumps. We call this the ”novelty” curve.

Our only job left now is to find the peaks in this ”novelty” curve, i.e. find the timestamps for the sudden jumps in energy, we perform so by a basic peak-finding algorithm. And there we have it, the note onsets, based on an energy approach.

2.1 Detection Function / ”Novelty” Curve

For finding the onsets, we employ an energy-based approach for computing the detection curve/”novelty” curve as discussed before.

2.1.1 Local Energy Function

Often, the beginning of a musical note coincides with a rapid rise in the energy of the signal. Leveraging this observation, a simple method to identify note onsets is to convert the signal into a local energy function that represents the energy level of the signal at each moment in time, and subsequently identify abrupt changes in this function.

Let $x : Z \to R$ be a signal. Furthermore, let $w : [-M, M] \to R$ for some $M \in N$ be a bell-shaped window function centered at time zero (e.g., a Hann window) (This window function is used to have better results in the next computational processes, i.e. to minimise the effect of large distanced signal points). The local energy of x with respect to w is defined to be the function $E_w^x : Z \to R$ given by

$$E_w^x(n) = \sum_{m=-M}^{M} |x(n+m)w(m)|^2 = \sum_{m \in Z} |x(m)w(m-n)|^2$$

2.1.2 Discrete Derivative and Half-Wave Rectification

To assess alterations in energy, we compute the derivative of the local energy function. In discrete scenarios, the most straightforward approach to accomplish this is by calculating the difference between two consecutive energy values. Moreover, given our focus on energy increments rather than decreases, we retain only the positive differences while assigning zero to the negative differences. The last step is known as half-wave rectification and is notated as:

$$|r|_{\geq 0} = \frac{r + |r|}{2}, r \in R$$

We thus obtain an energy based novelty function $\Delta_{Energy} : Z \to R$ given by:

$$\Delta_{Energy}(n) = |E_w^x(n+1) - E_w^x(n)|_{\geq 0}$$

We use this ”Novelty” curve, i.e. find the peaks, to get the onset timings.

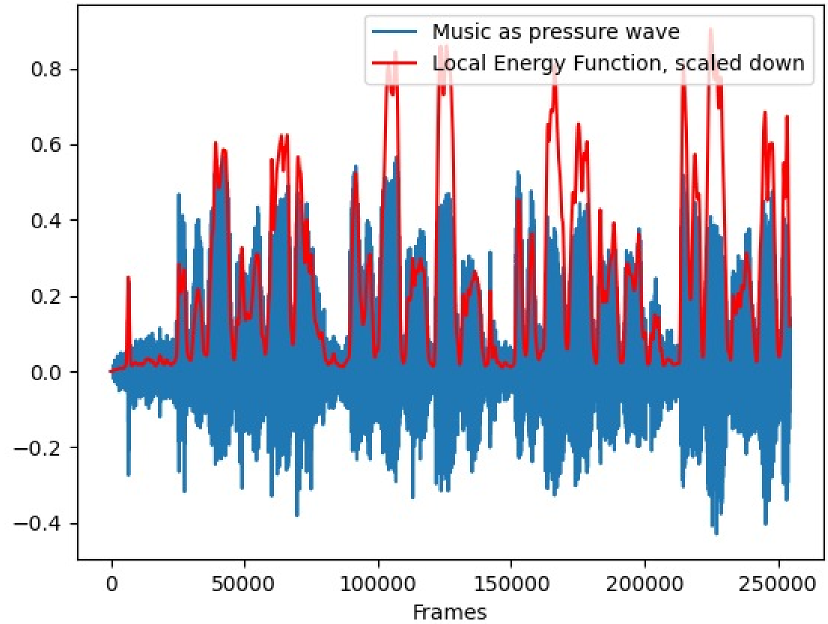

Figure 1: Pressure and Local Energy vs Time Frames

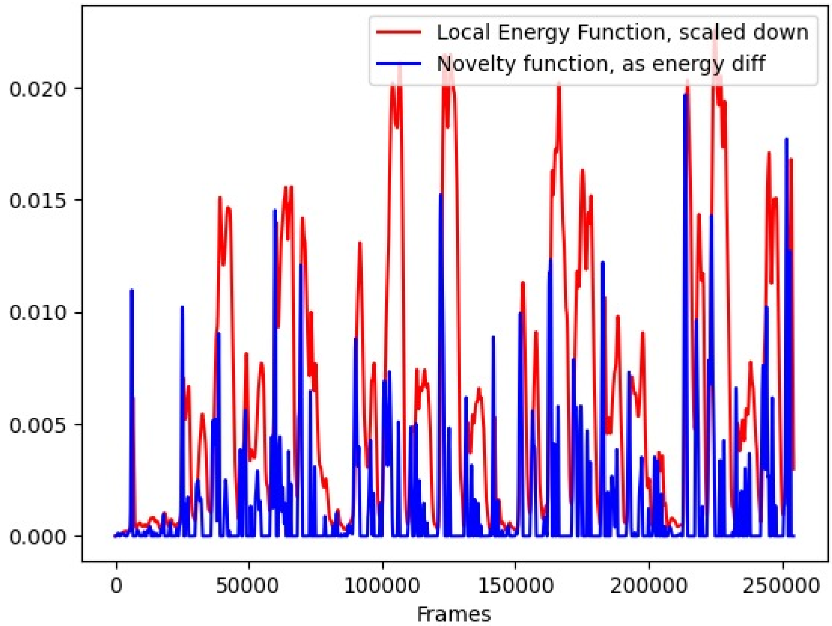

Figure 2: Local Energy and Detection Curve vs Time Frames

2.2 Peak Finding

Now, that we have the "Novelty" curve, the peaks of the same are the points of sudden energy bursts, i.e. note onsets. We find the peaks from our "Novelty" curve and get our Note onsets.

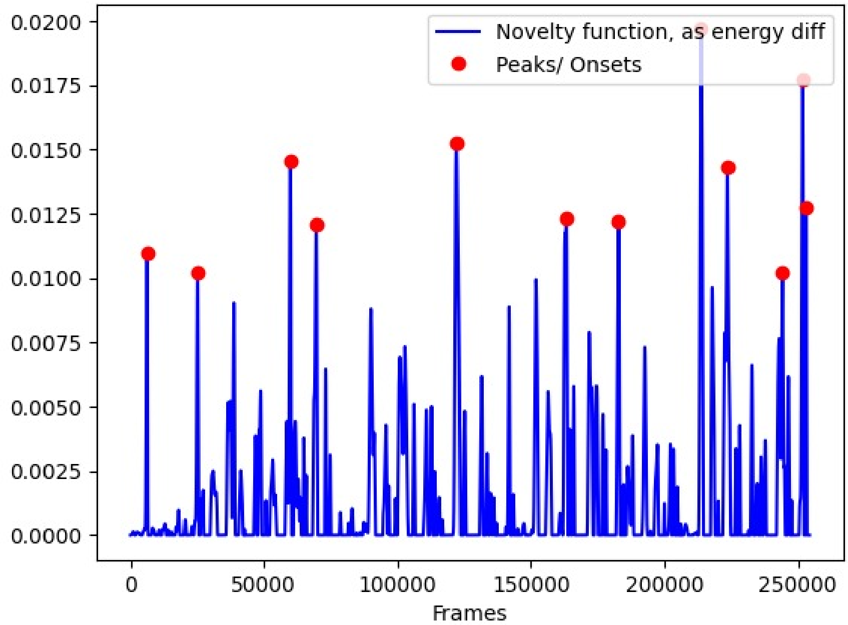

Figure 3: Novelty Curve with the peaks marked vs Time Frames

Here we have used a basic algorithm for peak finding which relies on neighbourhood values to find a local maxima (built-in in library scipy.signal).
An important thing to notice is that not all peaks are picked, the relatively small peaks are generally due to noise and are therefore ignored. This is done by the tuning parameter 'prominence' which determines which value should be considered a peak, even if it is the local maxima.
The peaks are the Note onsets required, which we understand, changes by changing the tuning parameter, a good approach for that is to check for the tuning parameter over a large number of note-onset datasets.

3 Conclusion

So, in conclusion, we have successfully demonstrated onset detection in music by finding peaks in a novelty curve, found through an energy-based approach.

4 Code

All the code has been done in Python. The code is given below. Plots shown here are based on a humming sound to be found here.

```
import sys
import numpy as np
import librosa
from scipy import signal
import matplotlib.pyplot as plt

x,sr = librosa.load(sys.argv[1] if len(sys.argv)>1
        else 'test.wav') # loading the audio data
x_duration = len(x)/sr
N = 2048 # Frames we want to consider in a window
w = signal.hann(N) # Choosing the Hann window

# computing the local energy function
x_sq = x**2
# The local energy formula given
energy_loc = np.convolve(x_sq,w**2,'same')
plt.plot(x)
plt.plot(energy_loc/25,'r-')
plt.show()

# Discrete time differentiation and half-wave rectification
energy_loc_diff = np.concatenate((np.diff(energy_loc),
                np.array([0,])))
# an ending zero for introducing the last 0-change
novelty_energy_func = np.copy(energy_loc_diff)
novelty_energy_func[energy_loc_diff < 0] = 0
# Half-wave rectification
plt.plot(range(len(novelty_energy_func)),
            novelty_energy_func,'b-')

# finding the peaks in the novelty function for onsets
peaks, props = signal.find_peaks(novelty_energy_func,
                prominence=0.012)
# finding the peaks, using scipy.signal,
# tuning param prominence set to 0.012
```

```
T_coef = np.arange(novelty_energy_func.shape[0])
peaks_sec = T_coef[peaks]
plt.plot(peaks_sec, novelty_energy_func[peaks], 'ro')
plt.show()

print('The onset times are : ')
print(*list(peaks_sec/sr))
```

5 References

- Tutorial on Onset Detection in Music Signals - Juan Pablo Bello, Laurent Daudet, Samer Abdallah, Chris Duxbury, Mike Davies, and Mark B. Sandler, Senior Member, IEEE
- Fundamentals of Music Processing Using Python and Jupyter Notebooks - Meinard Müller
- Onset Detection, Audiolabs tutorial
- Novelty curve based on Local Energy, Audiolabs tutorial