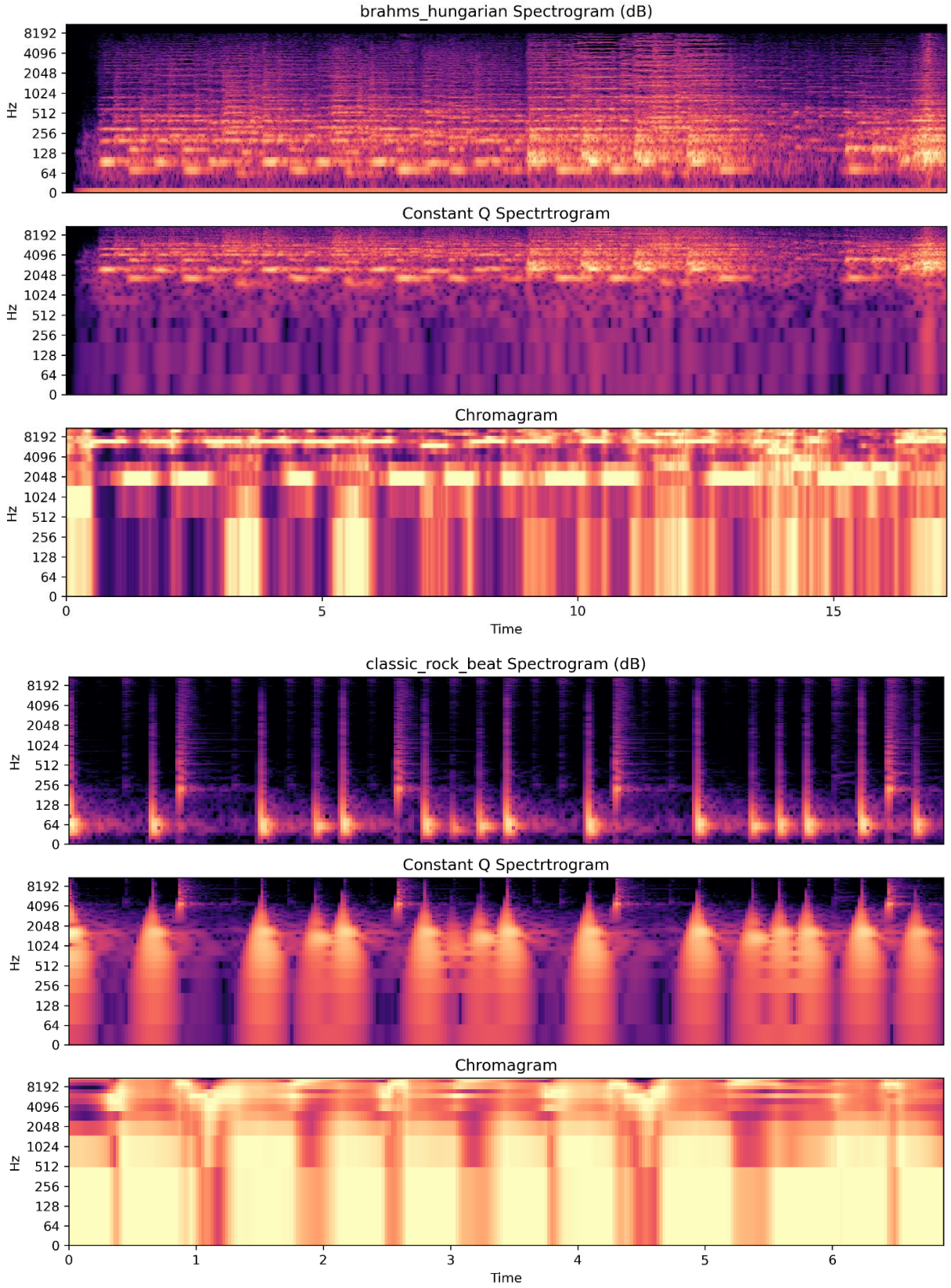
Audio Processing

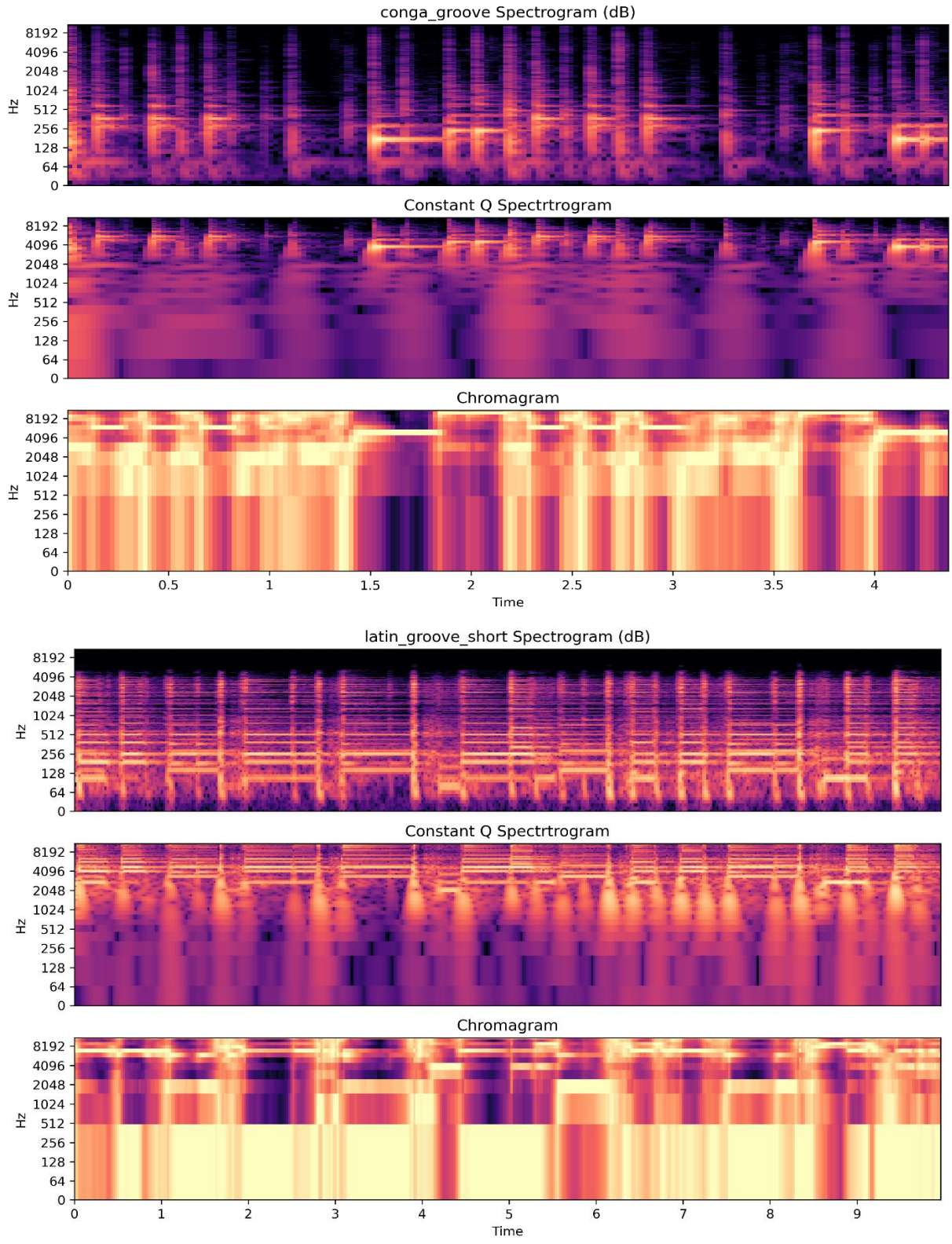
Ex06 – Report

Student: Anh Huy Bui

ID: 293257

Problem 1:

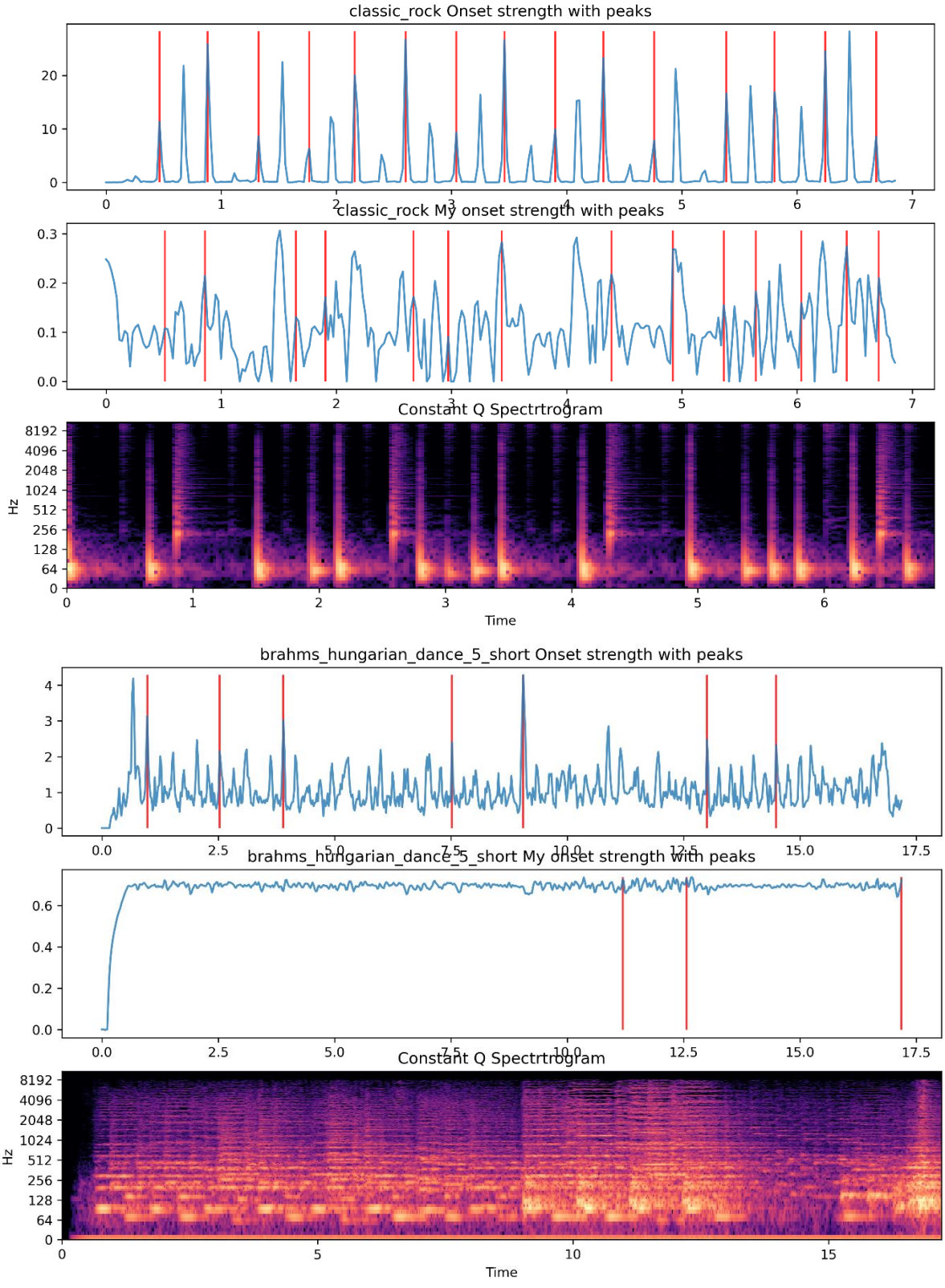

brahms_hungarian Spectrogram (dB)
Constant Q Spectrtrogram
Chromagram
classic_rock_beat Spectrogram (dB)
Constant Q Spectrtrogram
Chromagram
Hz
8192
4096
2048
1024
512
256
128
64
0
Time
0
5
10
15
1
2
3
4
6

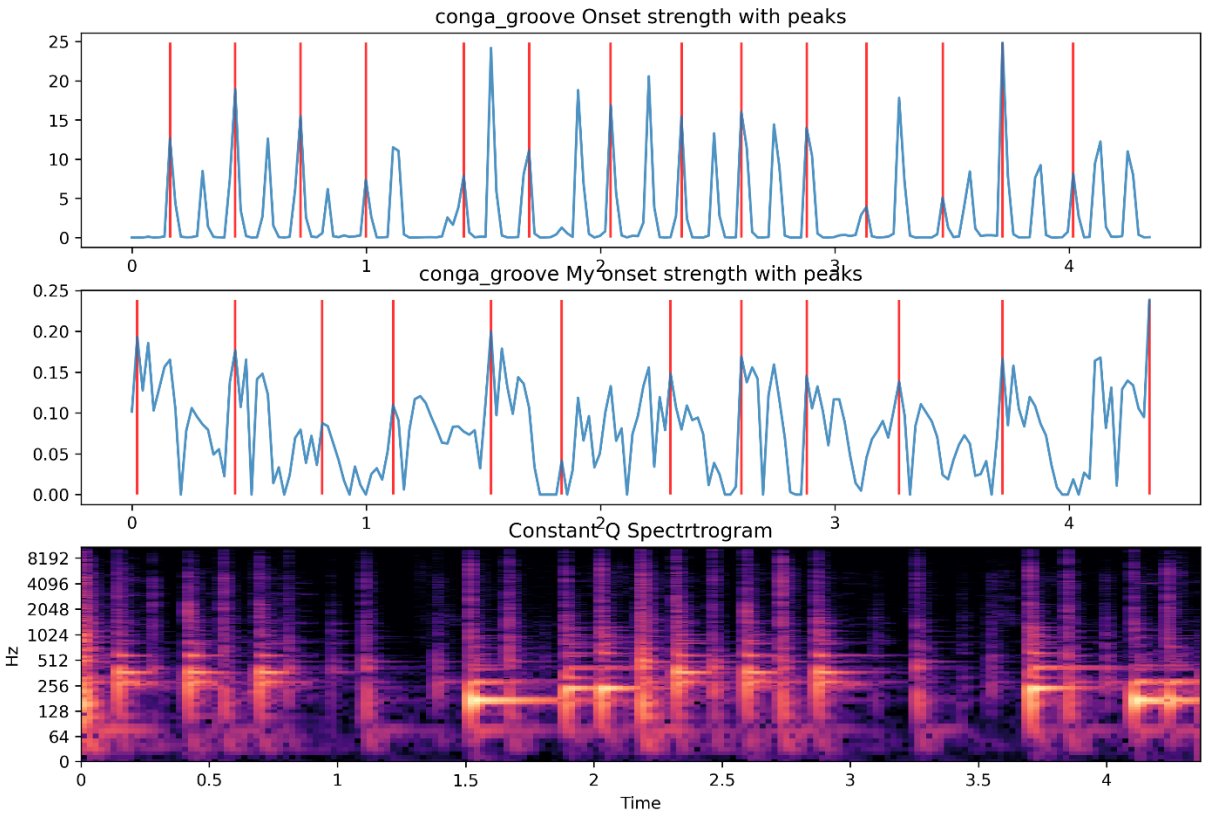


Comparing to STFT, constant Q spectrogram at the same sample (time frame) has a wider bottom and thinner top. This is because of resolution of time varies depends on frequency. This also happen in the horizontal axis, when the lower frequencies are thicker and higher frequencies are thinner.

Chromagram has nearly same color in all regions (from low to high frequencies). This is because its amplitude does not depend on the audio amplitude. Frequencies within a single time frame are distributed around a main/ major note.

Problem 2:


classic_rock Onset strength with peaks
classic_rock My onset strength with peaks
Constant Q Spectrtrogram
Hz
Time
brahms_hungarian_dance_5_short Onset strength with peaks
brahms_hungarian_dance_5_short My onset strength with peaks
Constant Q Spectrtrogram
Hz
Time



Researching on pick_peak, some parameters have to be taken into account: pre_max, post_max, post_avg, delta, wait.

1. pre_max, post_max determine range for calculating max value. Increase this may result in less frequent appearance of peaks, but more accurate picked ones.
2. Wait value determines minimum amount of time before the next peak is picked. Increase wait value will return less frequent appearance of picked peaks.
3. Delta value is a threshold offset for mean value of calculating range.

Bonus:

Comparing to the already-made function, my own implement seems to be less accurate, this is because the already-made version applied some filters to get a better result with less noise while keeping the signal from data loss.