Report

Problem 1-a

Code: onset_ energy.m

Problem 1-b

Code: synth_onset.m

Problem 1-c

Code: main.m, part 1-c

File: Pop_onset_energy.wav

Params: win = hamming(512); hop = 512; th = 0.11; frameLen = 512; frameHop = 512;

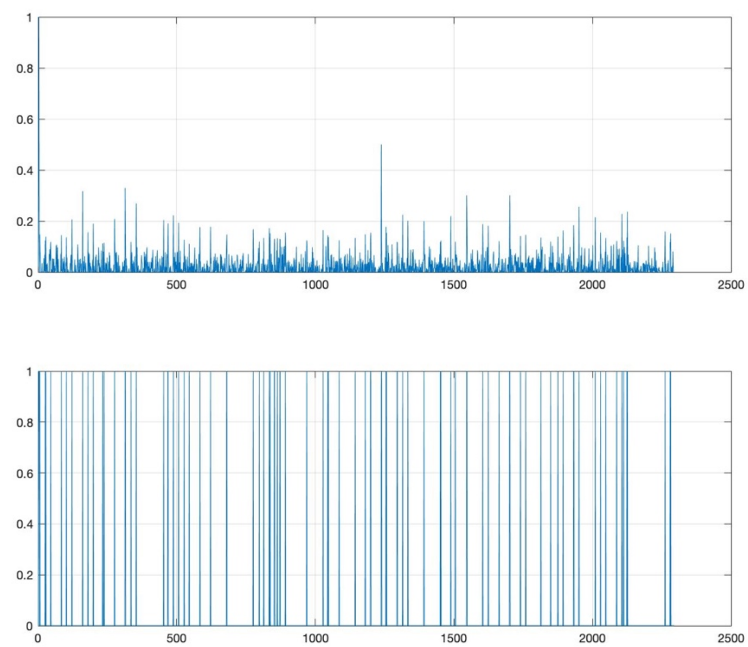

Figure 1 Energy-based onsets detection on Pop.wav

Because different frequency bands are not handled but treated equally, human voice change and beats are detected at the same time. Some soft onsets are not detected.

Problem 1-d

Code: onset_spectral.m

Problem 1-e

Code: main.m, part 1-e

File: Pop_onset_spectral.wav

Params:

win = hamming(512);hop = 512;th = 0.095;gamma = 1;frameLen = 512;frameHop = 512;

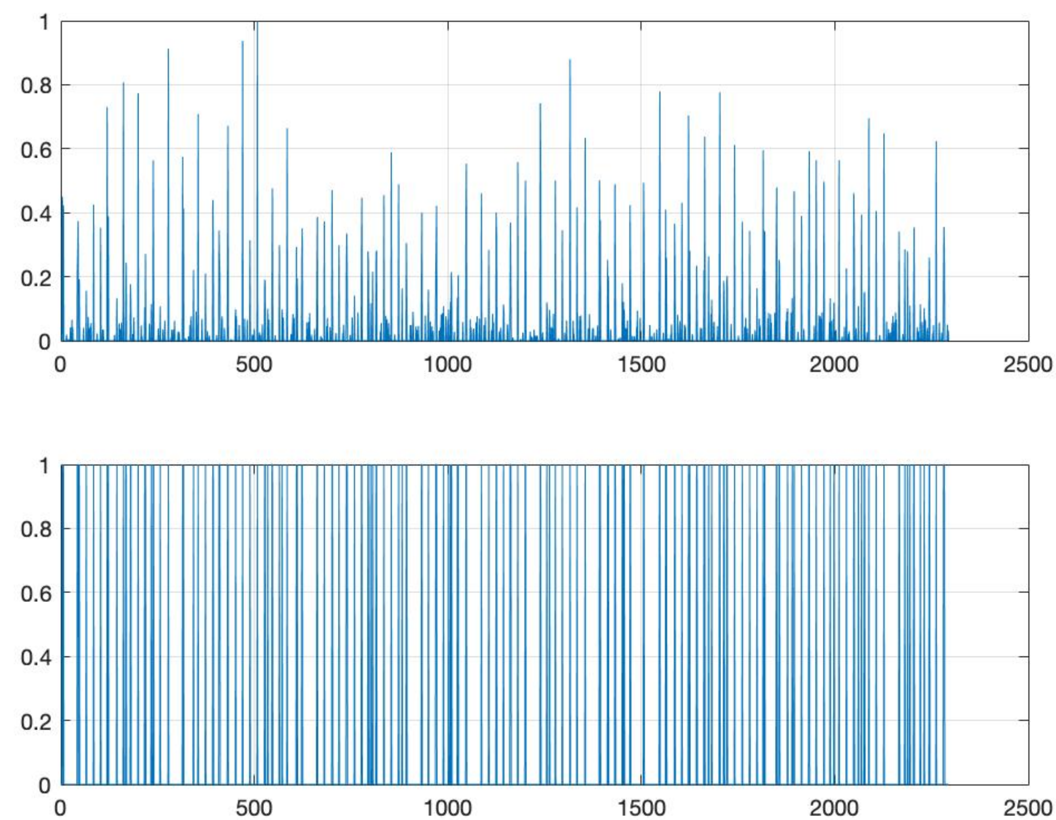

Figure 2 Spectral-based onsets detection on Pop.wav

Better performance on percussive beats detection when human voice and percussive onsets happens very closely. Which means this method can separate human voice change and percussive beats change, because it treats different frequencies differently. Better performance on softer onsets, as it can detect half a clap onset.

Problem 1-f

Code: main.m, part 1-f

File: Shumann_onset_energy.wav, Shumann_onset_spectral.wav, Haydn_onset_energy.wav, Haydn_onset_spectral.wav

Params(Shumann):

Energy: win = hamming(512); hop = 512; th = 0.069; frameLen = 512; frameHop = 512;

Spectral: win = hamming(512); hop = 512; th = 0.224; gamma = 1; frameLen = 512; frameHop = 512;

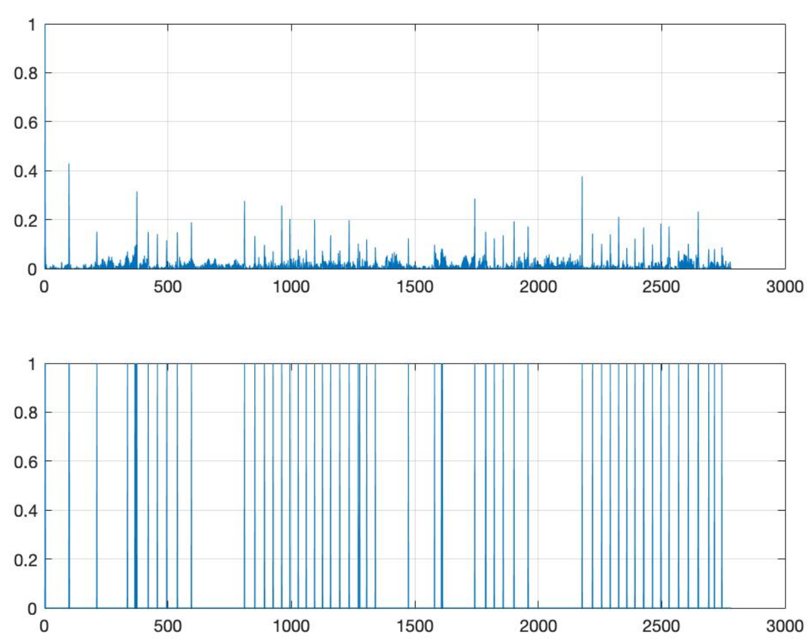

Figure 3 Energy-based onsets detection on Shumann.wav

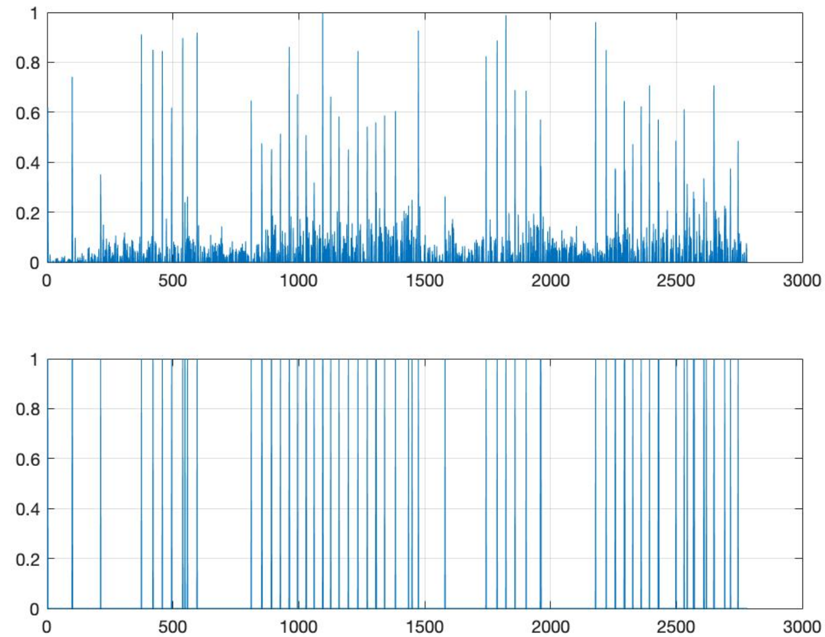

Figure 4 Spectral-based onsets detection on Shumann.wav

For the Shumann piece, although the beats are not strong and regular, the onsets of the piano notes are clearly stressed. So, both energy based and spectral based method work well.

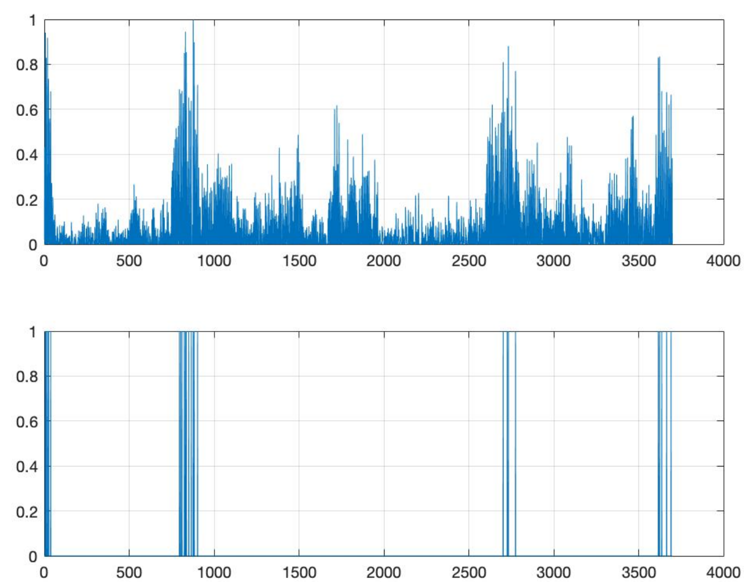

Figure 5 Energy-based onsets detection on Haydn.wav

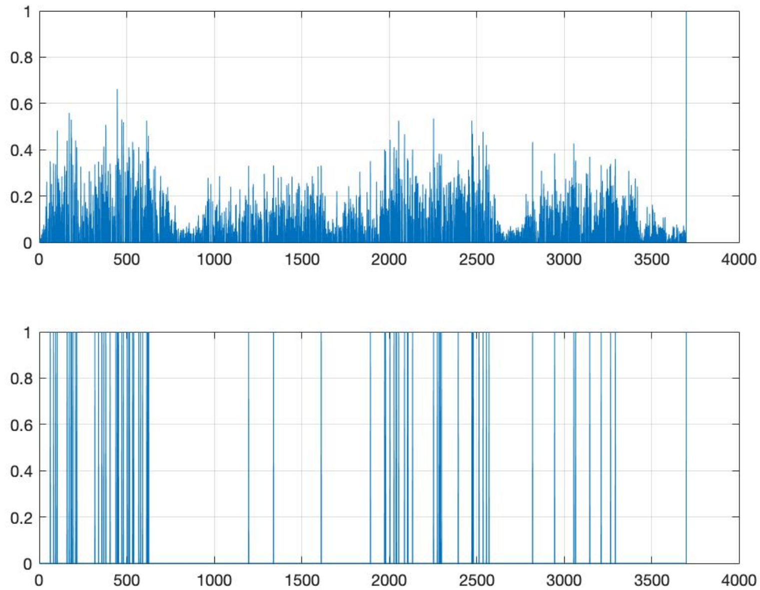

Figure 6 Spectral-based onsets detection on Haydn.wav

For the Haydn piece, both energy-based and spectral based method work poorly. The later one outperforms the former one as it has logarithmic compression stage which tunes gamma to enhance weak spectral components. There are several reasons for the dissatisfying result. For example, the music piece itself does not have strong onsets. Brass and string instrument sound starts very gently and accumulated softly which makes it hard to detect noticeable energy change. Besides, different instruments are played harmonically which lowered the spectral approach performance. Also, the thresholding is hard to set properly as seen in figure 6, the onsets energies are close to each other.

To improve this, we can separate the signal into different frequency bands during the preprocessing stage, and then combine results of each bands. Or use some kind of dynamic thresholding process.

Problem 2-a

Code: main.m, part 2-a

Problem 2-b

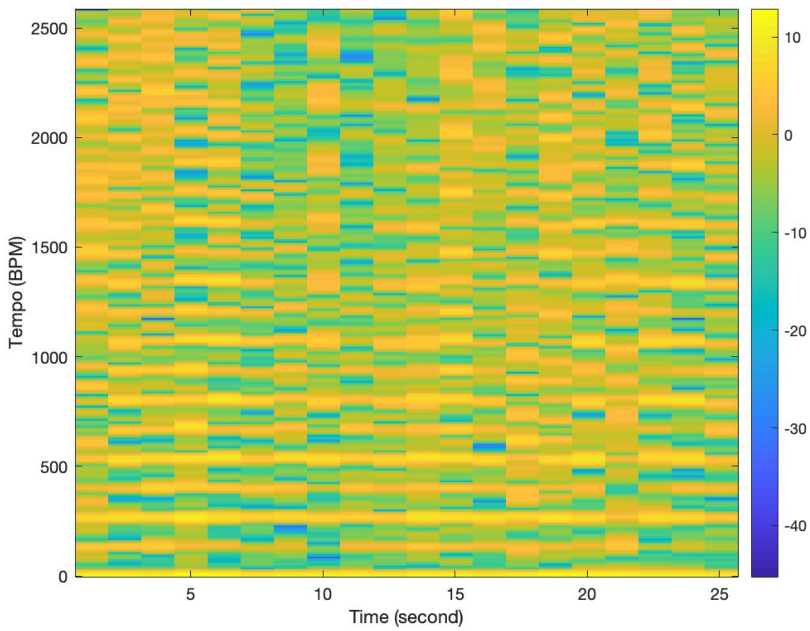

Figure 7 Fourier tempogram of Pop.wav

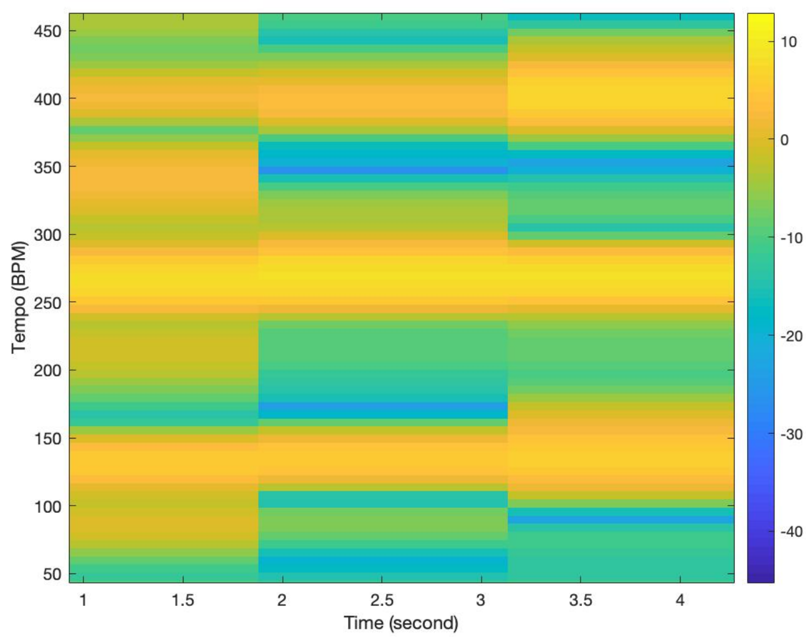

Figure 8 Zoom of Fourier tempogram of Pop.wav

We can see from the zoom picture of tempogram, there are several lightest frequency lines. The most predominant one is about 260 BPM, and the second is about 130 BPM, half of it. As we can hear from the audio or see from the onset strength curve, the onset detection methond treat down and up of a beat as two beats. So, the fundamental frequency, i.e. the tempo of the piece should be about 130 BPM.

Problem 3-a

Code: beat_dp.m

Problem 3-b

Code: main.m, part 3-b
File: Pop_beats.wav

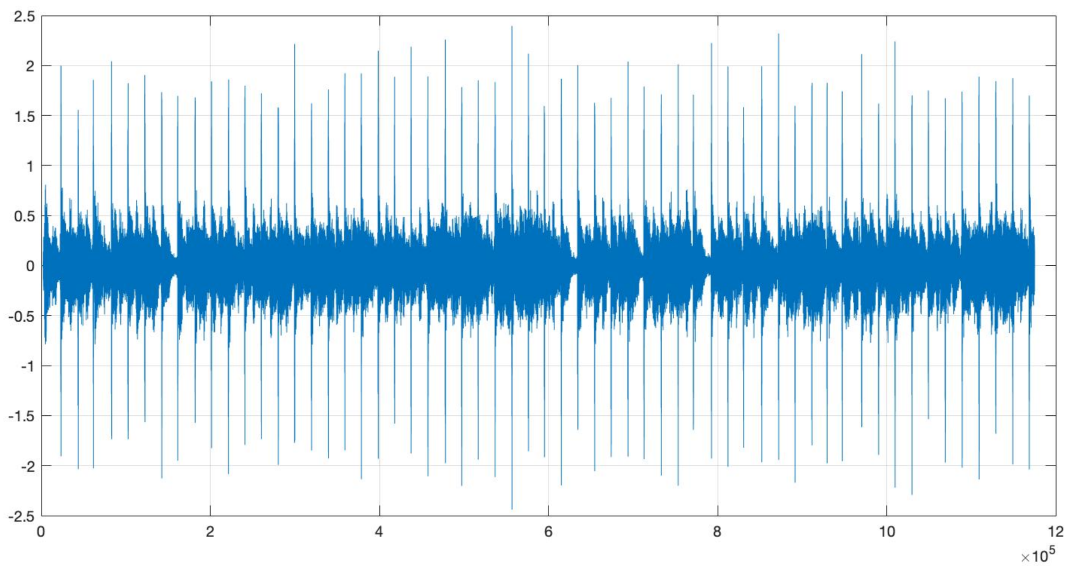

Figure 9 Beats tracking of Pop.wav

The beats tracking result of Pop.wav is great, just missed the first beat, and the others are accurate.