U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Assignment A3.

Pitch and Rhythm

Multimedia Systems

Alena Tesařová
Tsagkarakis Stylianos

June 5, 2020

Task 1 - Pitch

Instant Analysis

Depart from pitch_estimation.ipynb and for the 2 monophonic signals (mp_oboe.wav; mp_organ.wav), select a relevant frame (i.e. not silence), and plot the output of the autocorrelation function. Answer the additional questions:

1. By peak-picking (find the location of the maximum), what would be the estimation of the pitch/f0 in
    a. Hz
    b. MIDI
2. Is this value correct or incorrect? If incorrect, explain why this algorithm didn't work.

1. Oboe

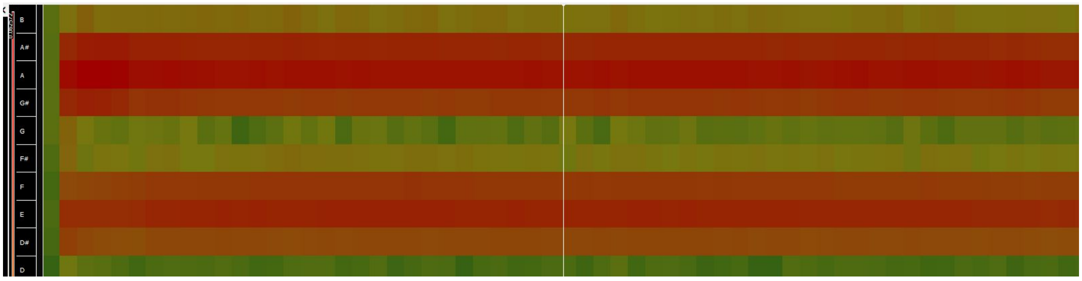
By analyzing the sounds maximum we generated the pitch in Hertz (441 Hz) and then we related that with the note by converting MIDI to Hz. Finally the value corresponded to A4.

1. Organ

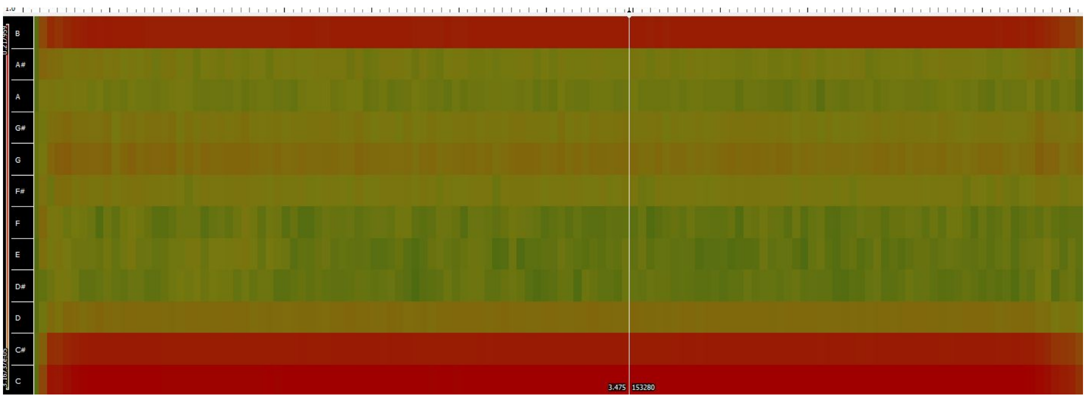
Using the same method as above we reached to 261 Hz and the note C4.

2. Correct values

To verify the results, we opened to sonic visuliser and we analyzed the files. Here are results:



Chromagram for oboe shows notes A + A# which corresponds with computed notes.



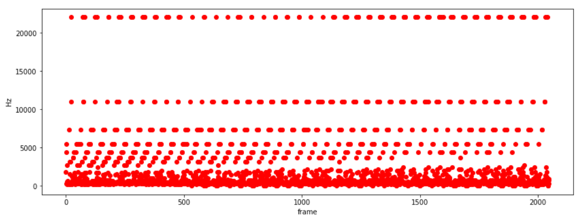
Chromagram for oboe shows notes B+C+C# which also corresponds with computed notes.

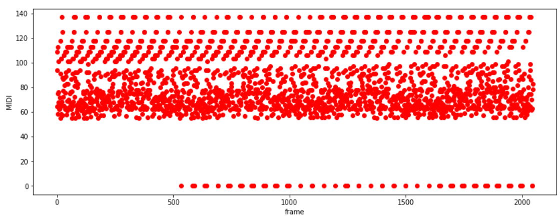
Global Analysis

On the previous task, your pitch estimates were frame-based. For the same signals, extend this frame analysis to the full sound, and for each sound:

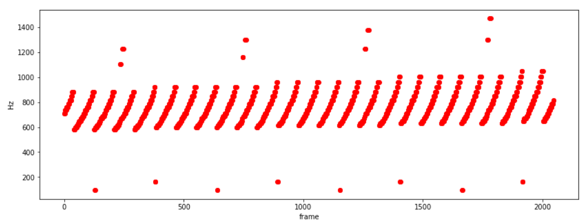
1. Obtain a plot of the pitch estimation; on the x-axis you'll have the frame number, on the y-axis the pitch/f0 in Hz.
2. Obtain a plot of the pitch estimation; on the x-axis you'll have the frame number, on the y-axis the pitch/f0 in MIDI Note (ex. A0, B0, B1)
3. Is any of these representations (1) or 2)) a chromagram? Explain your answer.

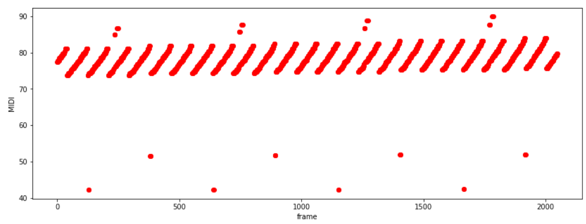
1. Plot #1 and #2 for oboe





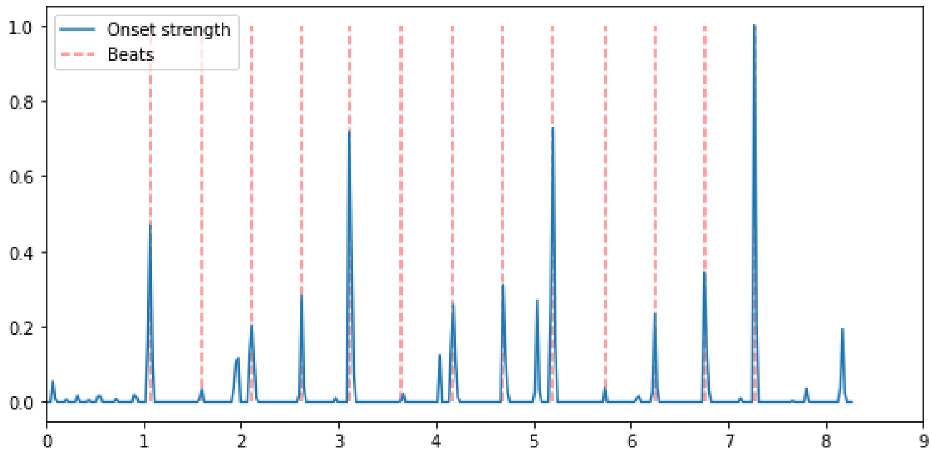
1. Plot #1 and #2 for organ





Task 2 - Rhythm

1. Do a subjective evaluation on the quality of the beat tracking.
2. This rhythm shows a steady constant tempo. Obtain the global tempo estimative in bpm (you can do it by hand, no need to code it).
3. How many downbeats are in this musical example? At what times?
4. What is the novelty function that you used? (i.e., the library implements). What functions do you know that could be used for the same end?
5. What would you obtain if you applied the autocorrelation function to this novelty function? Explain your answer.



1. ## Subjective Evaluation

   We think the quality is very good. We don't hear any deviation.

2. ## Our estimation:

   BMP = Beats per minute

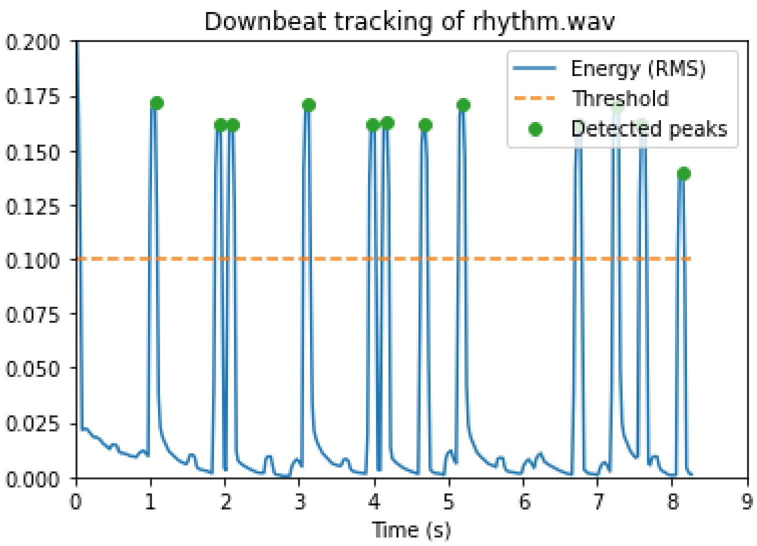
   13 clicks in 6.7 sec => 117 BPM
   60 / 6.65 * 13 = 117.29

3. ## Downbeats

   Due to the “loudness” of the beat we propose that there are 4 downbeats (values are approximately):
   - 0:01
   - 0:03
   - 0:05
   - 0:07

4. ## Novelty function

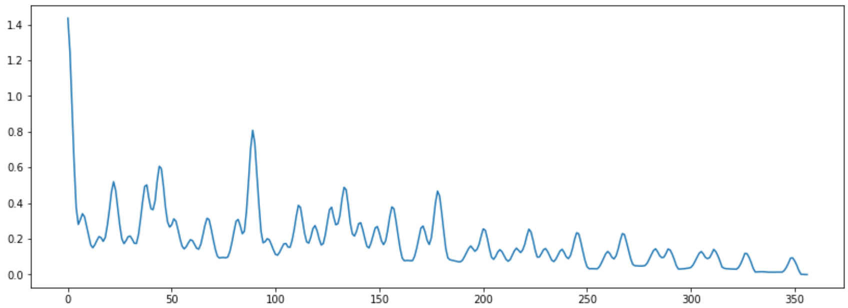
   The novelty function that we used is: librosa.feature.rms(), librosa.beat.beat_track(), scipy.signal.find_peaks(). All of them after some modification and transformation through analysis of the data.
   Some other functions:
   - To obtain an energy novelty function, we can perform half-wave rectification i.e. any negative values are set to zero. Equivalently, we can apply the function max(0,x)
   - The human perception of sound intensity is logarithmic in nature. To account for this property, we can apply a logarithm function to the energy before taking the first-order difference. Because log(x) diverges as x approaches zero, a common alternative is to use log(1+λx). This function equals zero when x is zero, but it behaves like log(λx) when λx is large.This operation is sometimes called logarithmic compression



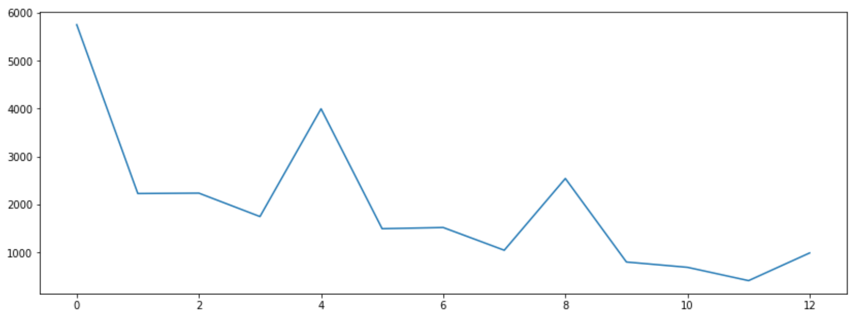
Our results are presented above.

5. Autocorrelation function to this novelty function

Oboe:



Organ:



Task 3 – Pitch and Rhythm I

Download voice.wav. This music excerpt contains 9 notes (part of a musical scale), singed by a feminine voice. With the help of Sonic Visualiser, answer the following:

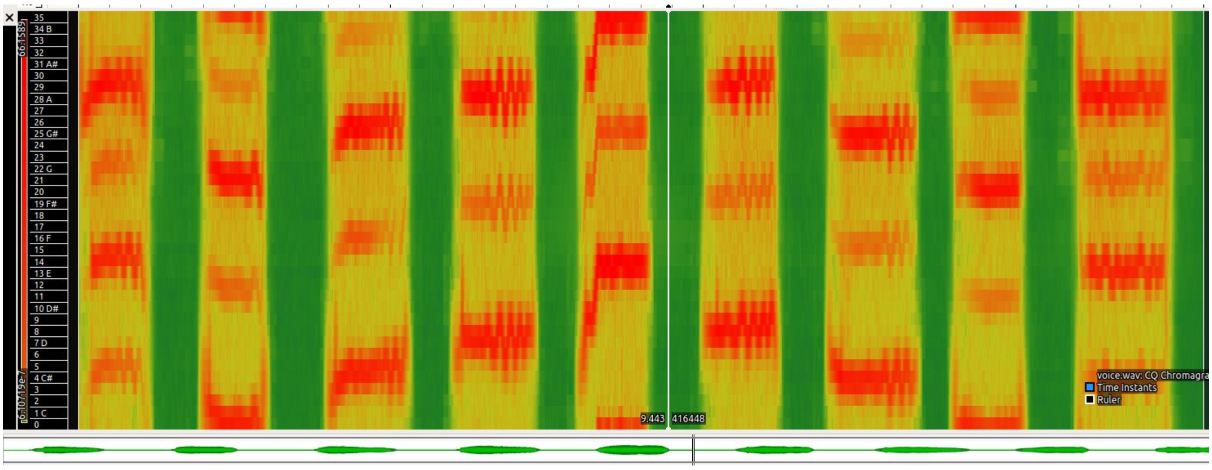
1. What is the chroma and height of the 9 notes?
2. Assuming each note represents a beat, what is the tempo (in bpm) of this excerpt.

Explain your answers and include the visualisations that best justify your rationale.

1. Chroma and height

As we can see from the chromagram below it is too complicated to understand clearly what are the notes.
Each note is represented by the **most dominant color** in the chromagram. Based on our intuition we chose below:

1. A or A#
2. C
3. G#
4. A or A#
5. E
6. (A or A#) or (D or D#)
7. G#
8. F#
9. E



2. Tempo (in bpm) of this excerpt.

It has 8 seconds and there are 9 notes => 67,5 BPM

Task 4 – Pitch and Rhythm II

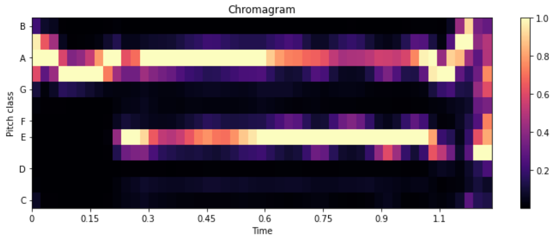
1. Choose the frames in which you have to estimate the pitch (exclude the silences).
2. Choose a way to average the estimates for each note across the different frames.

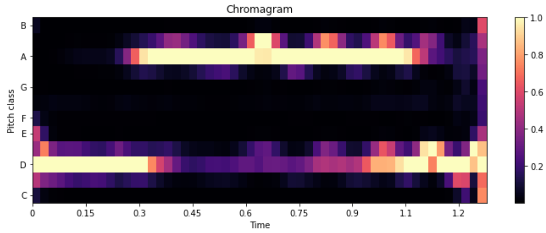
1. Choosing frames

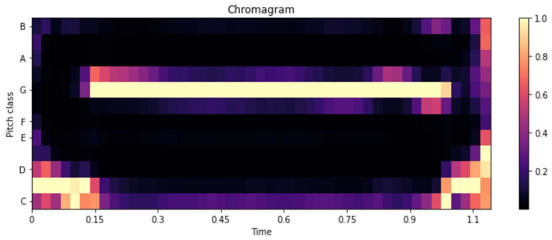
We chose the frames by doing some simple analysis:

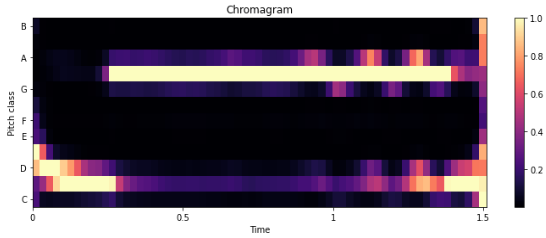
- We set a threshold of 50 db loudness
- We looped through the samples and discarded the ones that were below the threshold with the help of the function: *librosa.effects.split()*
- We split the sound in 9 parts based on that.

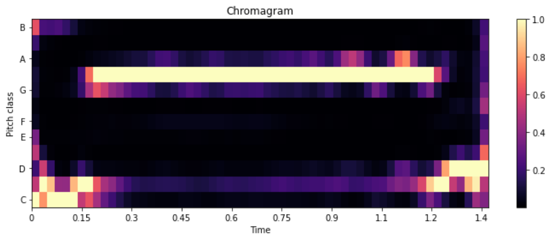
Generating chromagrams:

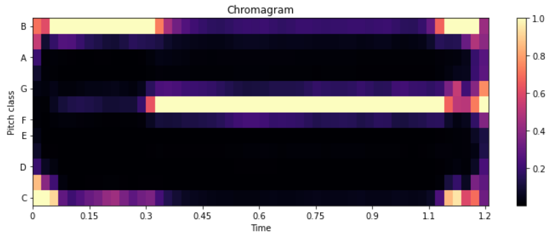


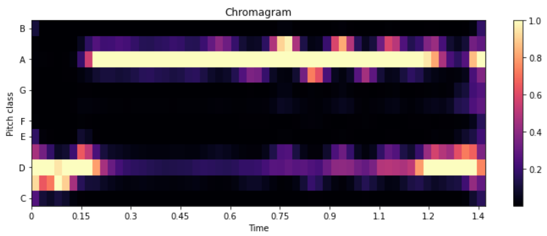


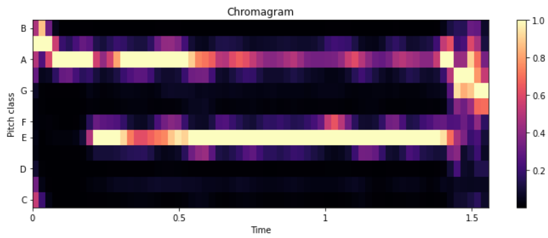


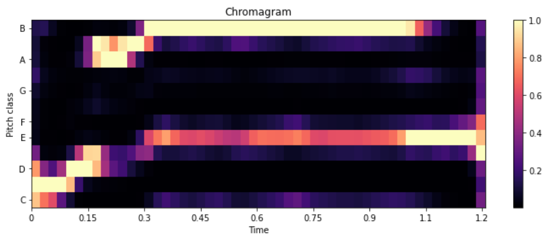












Analyzing chromagrams:

1. Note: **A**
2. Note: **G**
3. Note: **G#**
4. Note: **A**
5. Note: **B**
6. Note: **A**
7. Note: **G#**
8. Note: **F#**
9. Note: **E**

As we can see there are really similar to our predictions.