Digital Signal Processing

By Zhi Jie Huang

1. Introduction:

This is the second lab on Digital Signal Processing, the purpose of this lab is to explore the methods for content-based music information retrieval based on the MFCC coefficient calculated from the Lab1. In this lab, we will get a chance to explore the musical features such as rhythm and tonality.

1. Calculation of similarity matrix:

In data analysis, similarity matrix is a graphical representation of similar sequence in data series. The similarity matrix measures the cosine of the angle between the spectral signatures of frames i and j from the mfcc matrix. If frame i and j is very similar, the similarity matrix value should be very close to 1(i.e very close to red); if frame I and j is not similar, the similarity matrix value should be very close to 0(i.e very close to blue).

As you can see, the two classical is mostly blue and green, that means the classical music doesn’t have a lot of similar note been play during different note. As we know, classical music has wide spread of notes been played throughout the song. So it physically make sense.

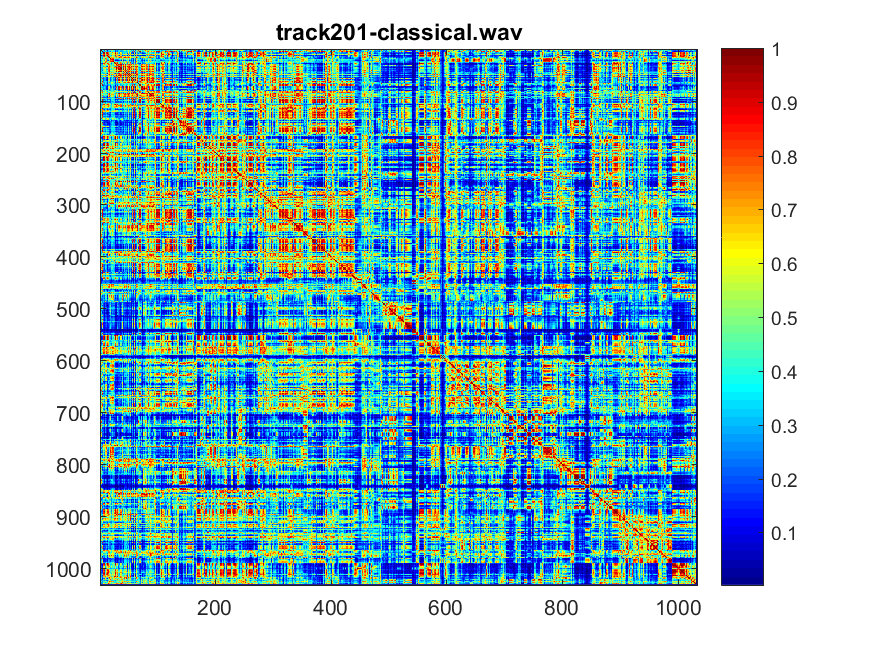
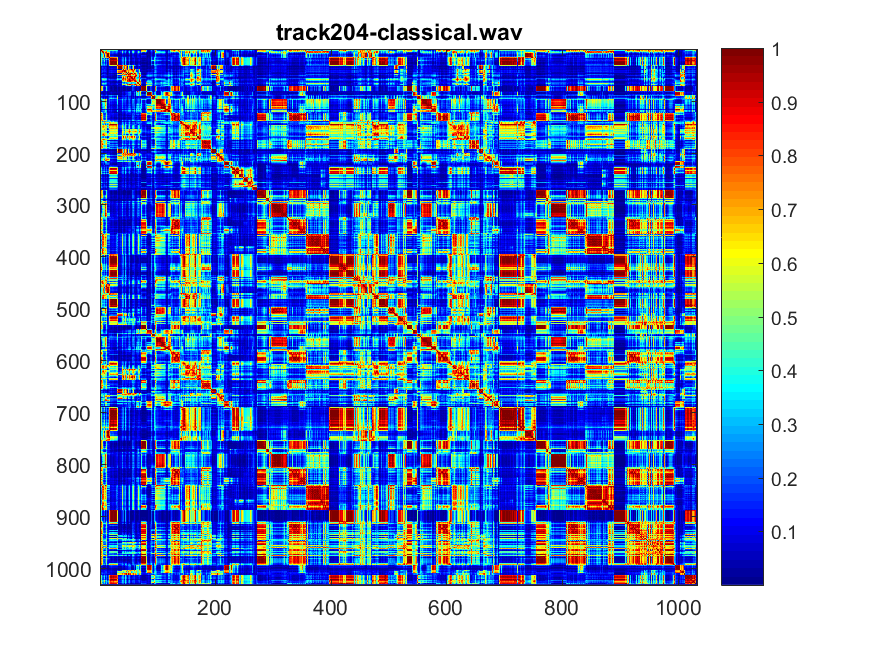


Fig 1: Similarity matrix for track201 Fig 2: Similarity matrix for track 204

Comparing the two electronics music, if you listen to track-370 electronic, you can feels a lot of different notes has been played through out the song, there aren’t a lot of repetitive note been played, therefore the similarity matrix is mostly blue. However, if you listen to track 396-electronics, you can tell there are multiple notes being play repetitive throughout different frames. Therefore, there are a lot of values closer to 1(red).

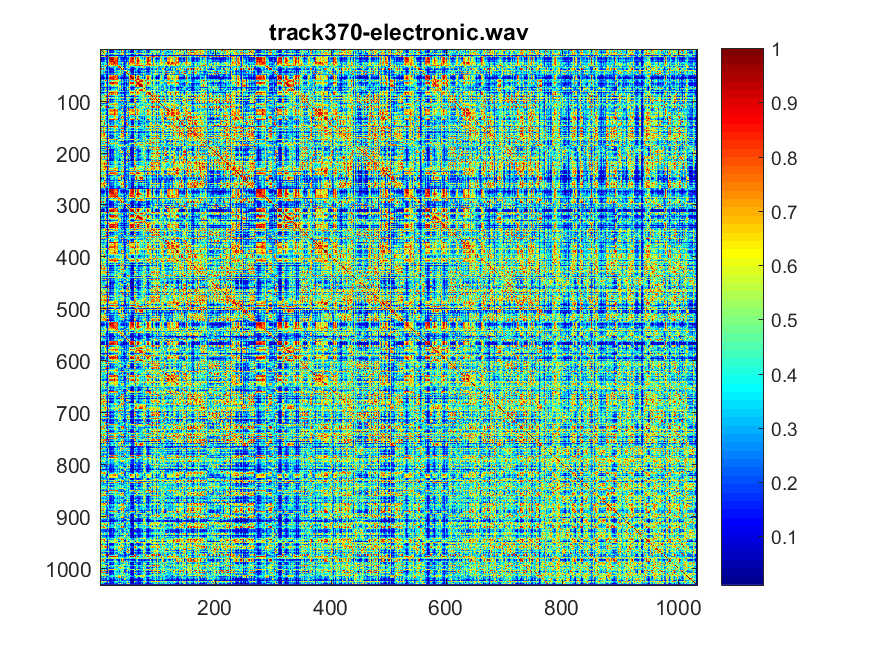


Fig 3: Similarity matrix for track370 Fig 4: Similarity matrix for track 396

Jazz music encompasses a wide range of music spanning a period of over 100 years, so it does have some similar note played throughout song for both of the track. So the frames are very similar between the i frame and the j frame.

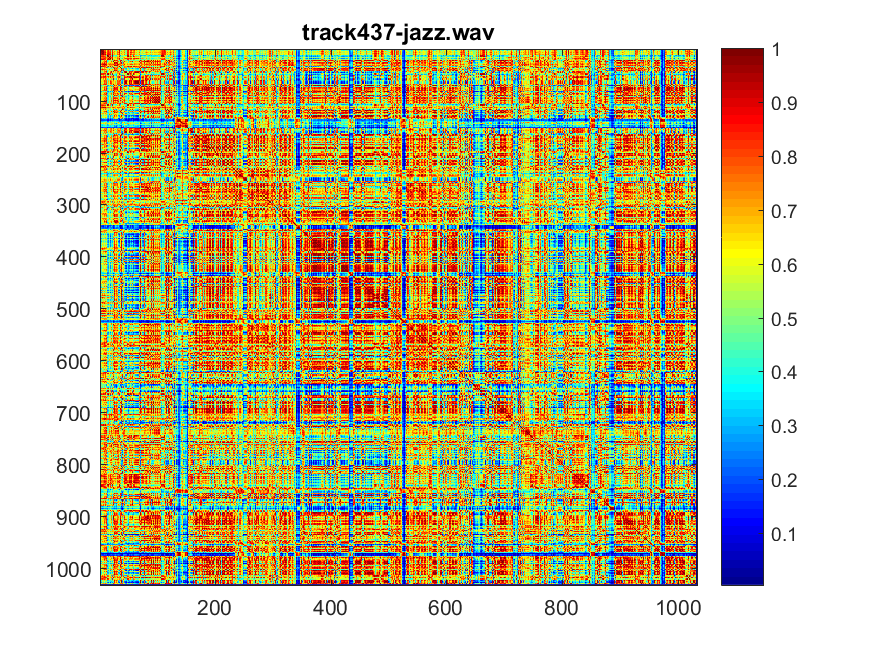
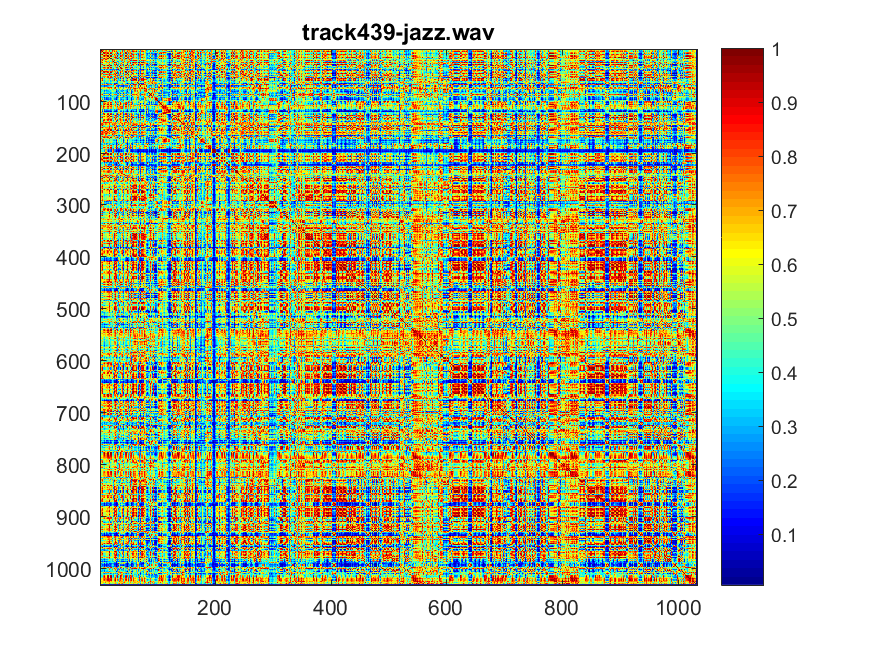


Fig 5: Similarity matrix track 437 Fig 6: Similarity matrix track 439

If you listen to track 463, there is a certain musical rhythm going on throughout the song, therefore, you can see notes been play every other 100 frames. However, for track 492 metal, there are too many notes been played repetitively, so it’s hard to distinguish a pattern for the song. The similarity matrix value for track 492 are all very close to each other.

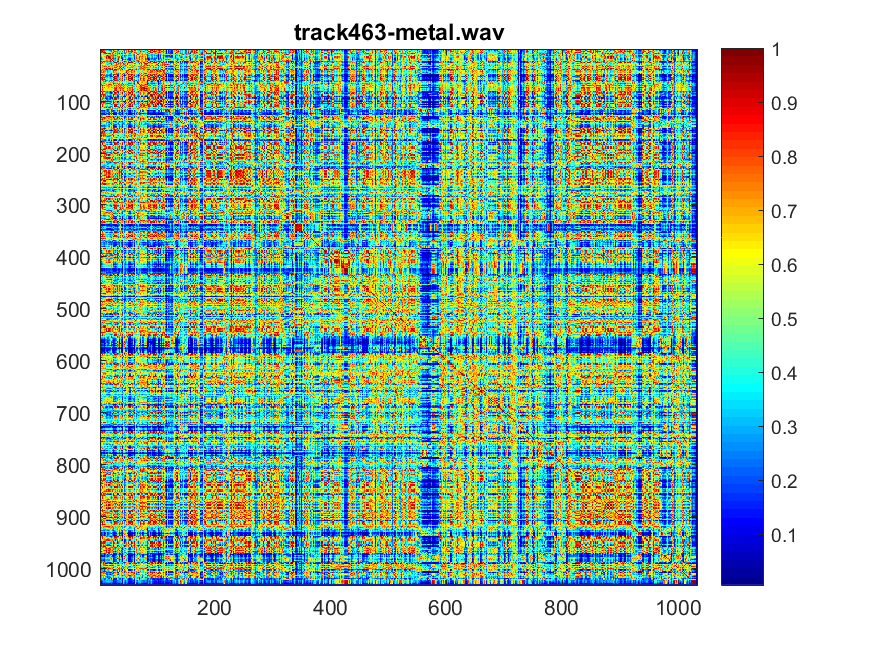
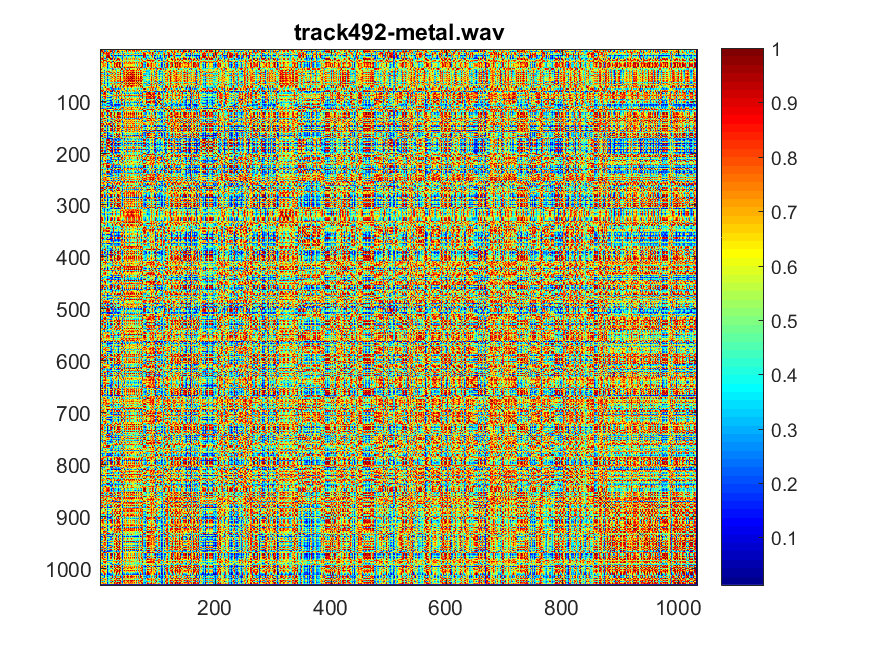


Fig 7: Similarity matrix track 463 Fig 8: Similarity matrix track 492

At the beginning of the track 547 rock, there are a lot of similar note been play. You can hear it from the middle of the song and see it in figure 9. For track 550-rock, while I was listening to the track, the notes been played doesn’t seems to have a repetitive pattern, therefore you can see from figure 10, the value for the similarity matrix is around 0.4~0.6.

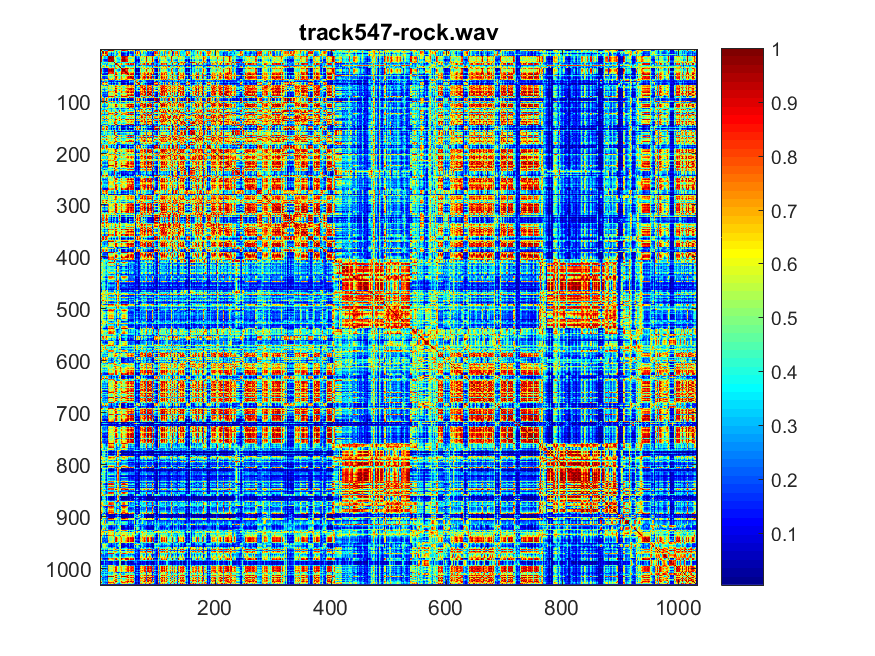
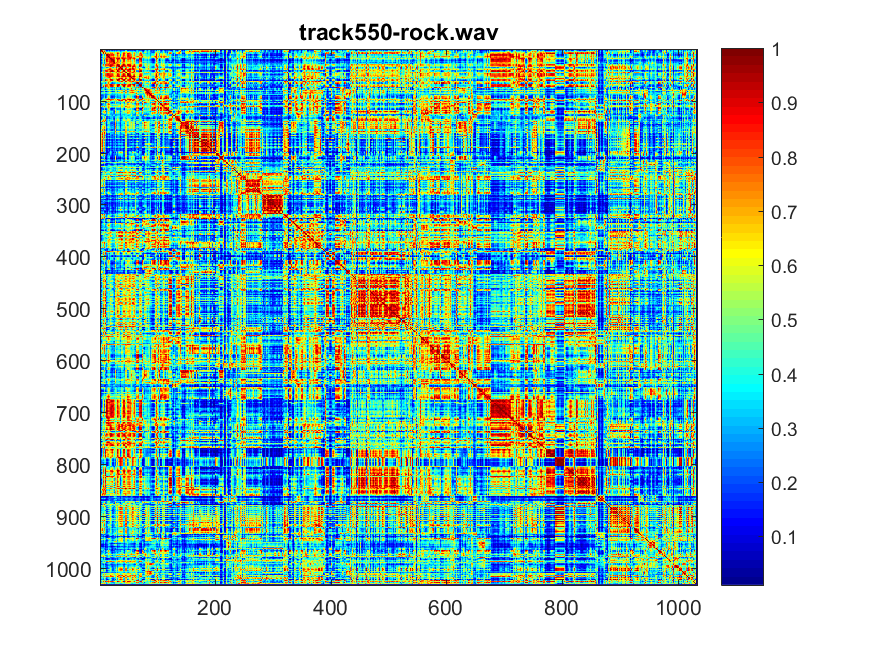


Fig 9: Similarity matrix track 547 Fig 10: Similarity matrix track 550

When you are listening to track 707, it is really quiet, and it doesn’t have a lot of similar note been play，except at frame 200, there was a flute playing for a little, it doesn’t have a lot of variation of note, therefore, you can see a lot of red in the middle. When you are listening to the track 729, it have very monotonous, it doesn’t have a lot of similar note play thoughout the song.

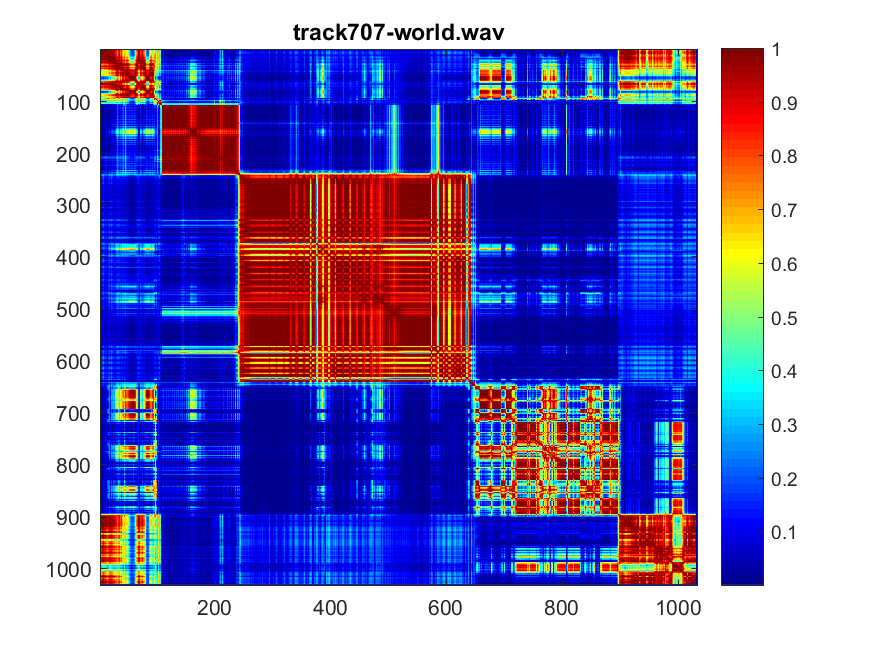
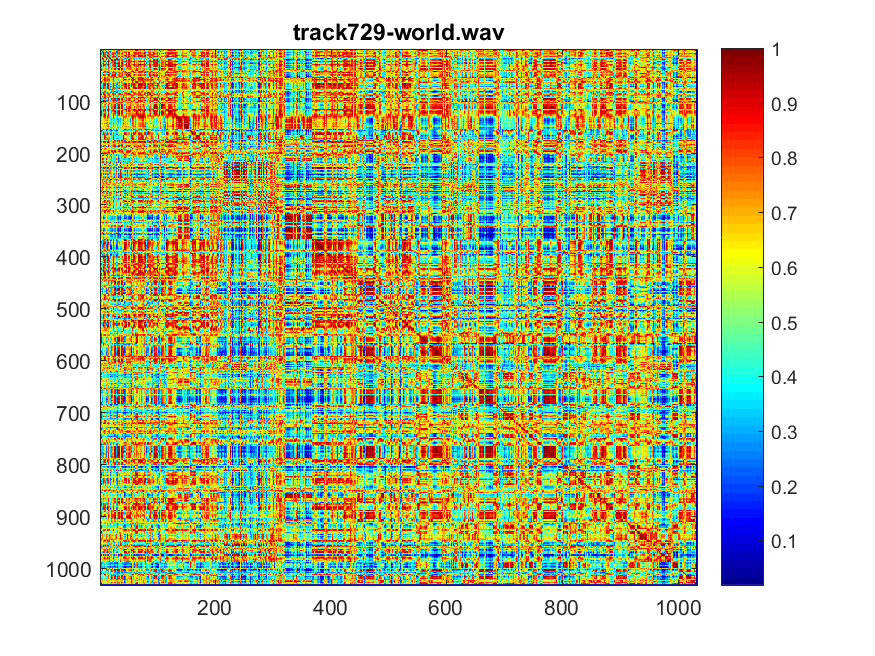


Fig 11: Similarity matrix track 707 Fig 12: Similarity matrix track 729

1. A first estimate of the rhythm

In this part of the lab, we are trying to use the similarity matrix information we calculated from part 2 to detect the presence of patterns that repeated every l frames away. The B(l) is the sum of all the entries on the lth diagonal.

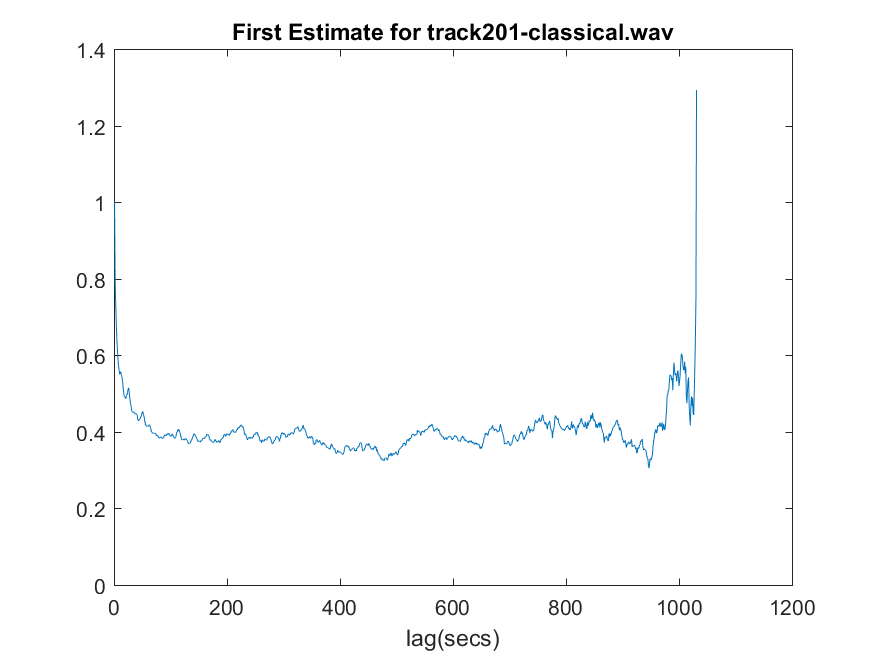
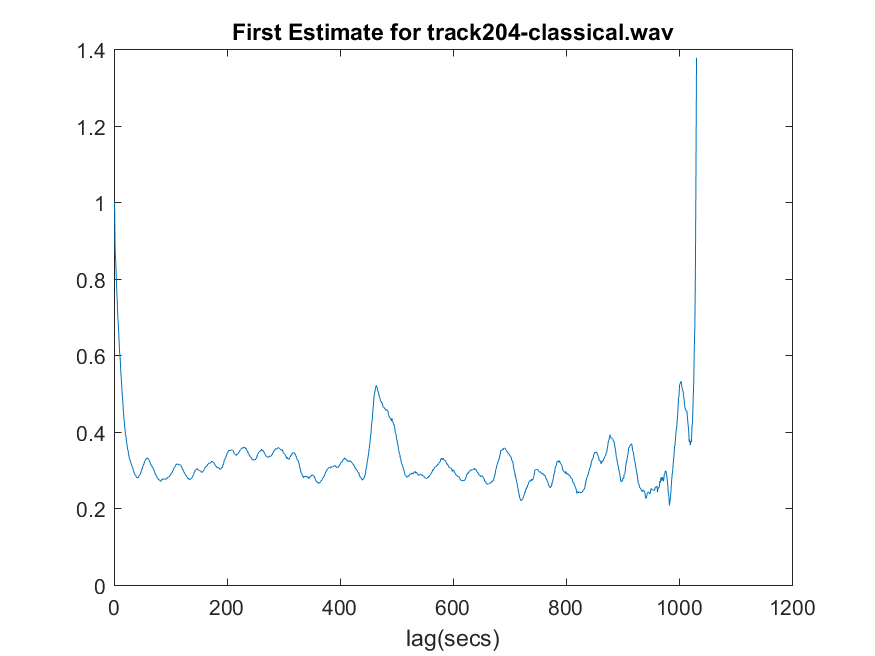


Fig 13:first estimate for track 201 Fig 14:first estimate for track 204

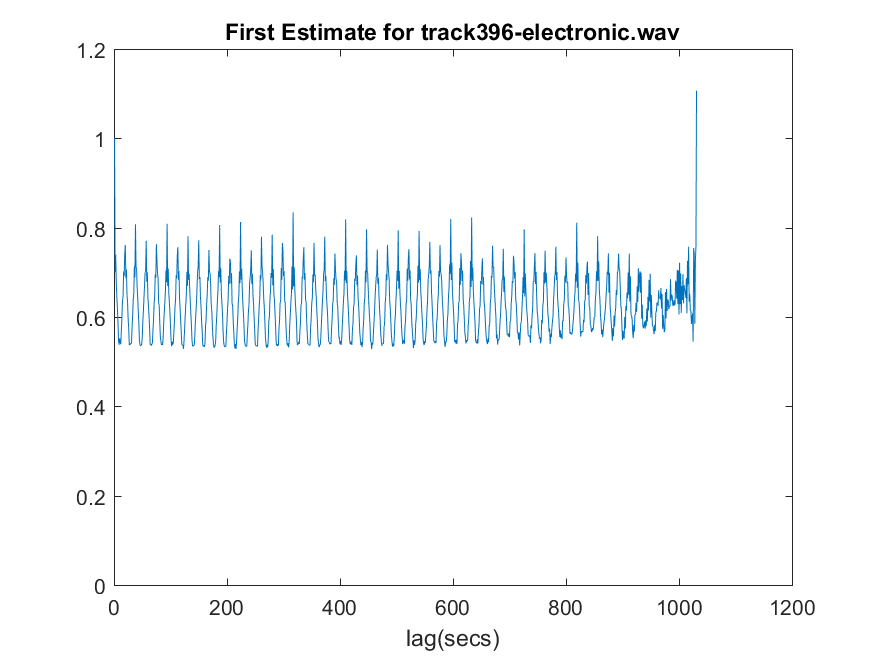
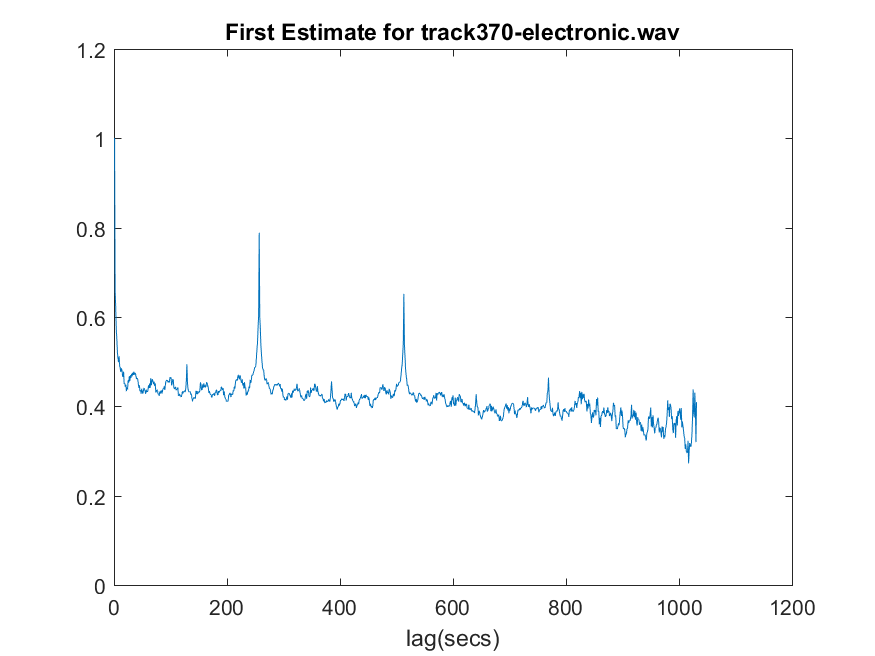


Fig 15: first estimate for track 370 Fig 16:first estimate for track 396

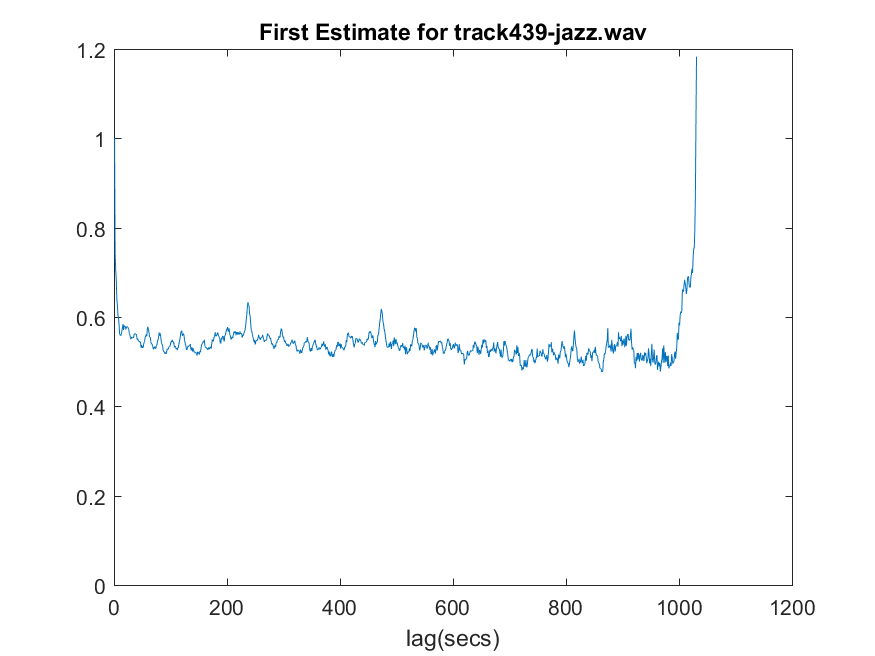
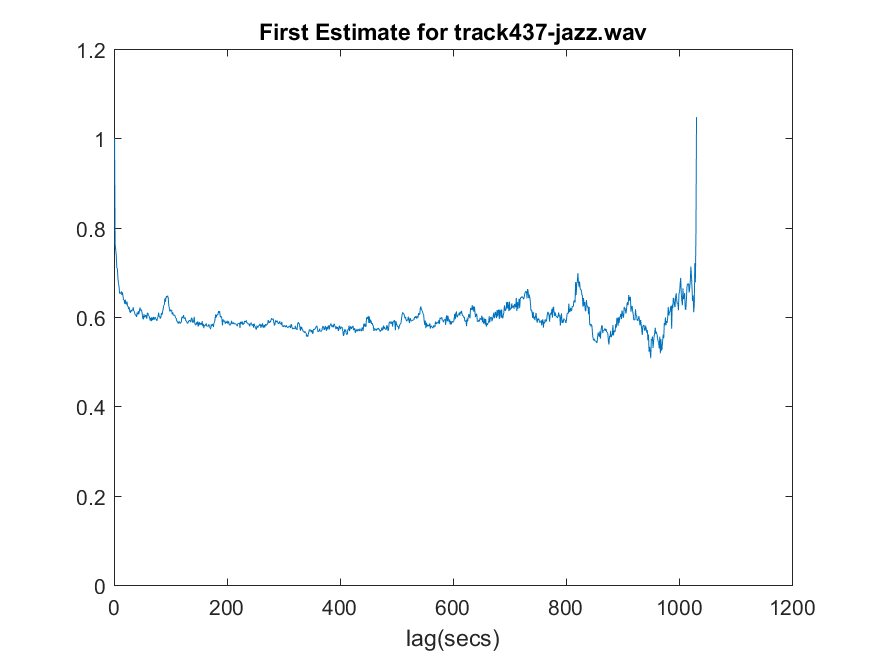


Fig 17: first estimate for track 437 Fig 18:first estimate for track 439

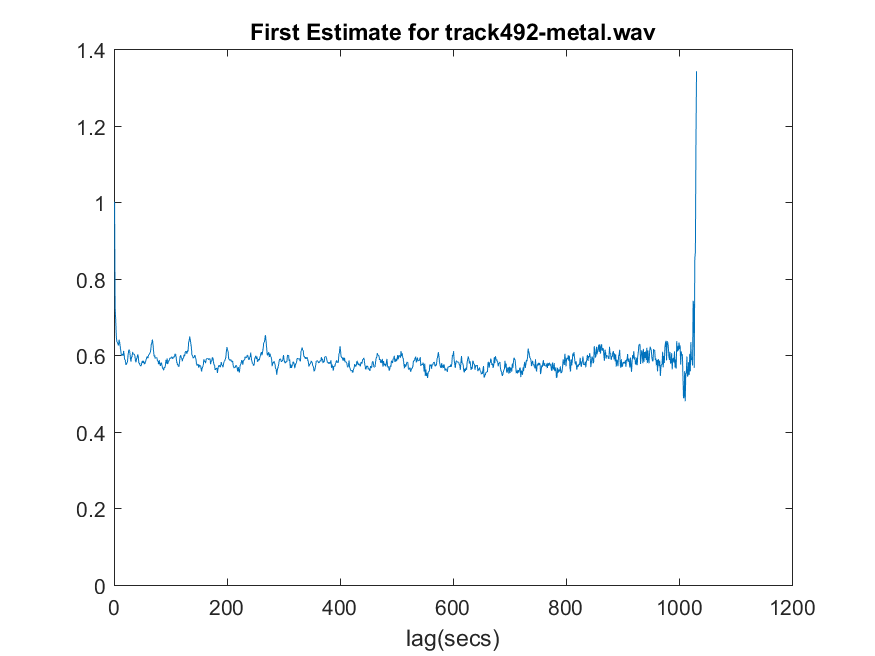
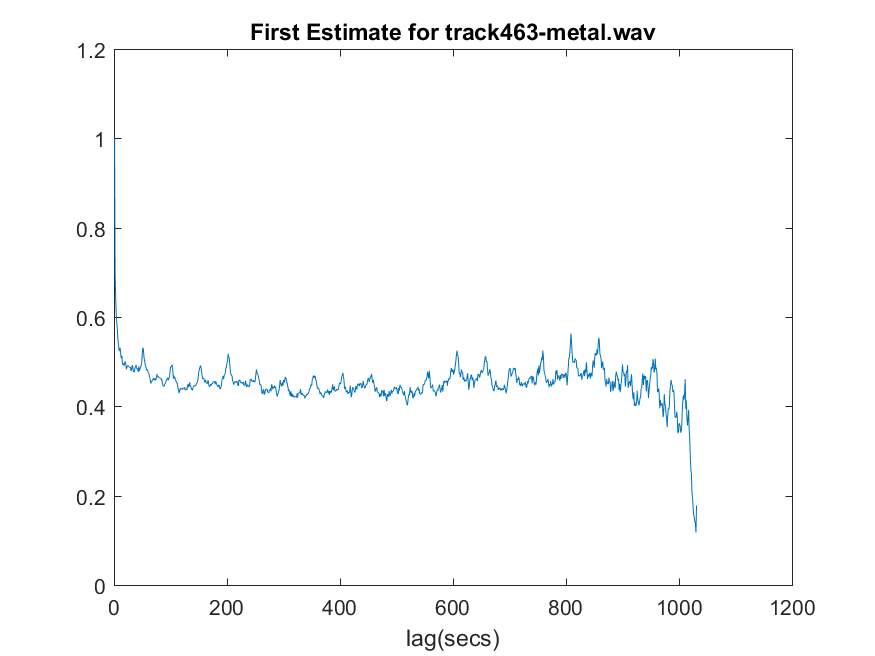


Fig 19: first estimate for track 463 Fig 20:first estimate for track 492

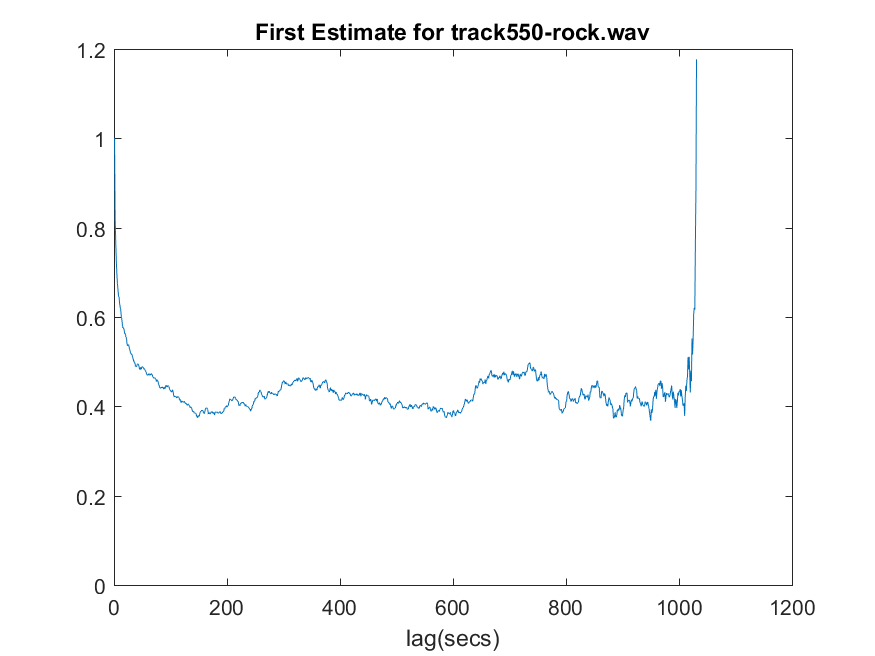
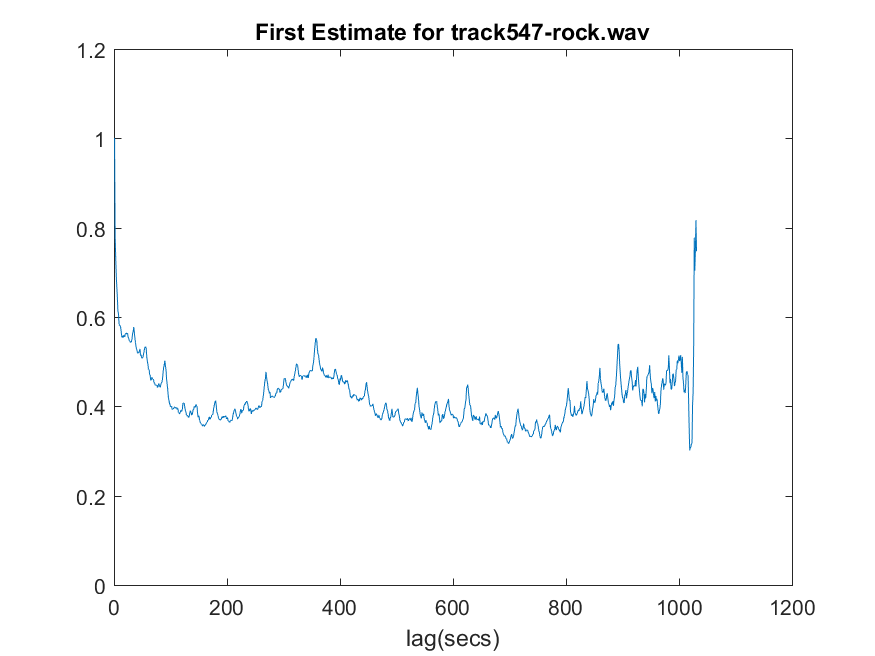


Fig 21: first estimate for track 547 Fig 22: first estimate for track 550

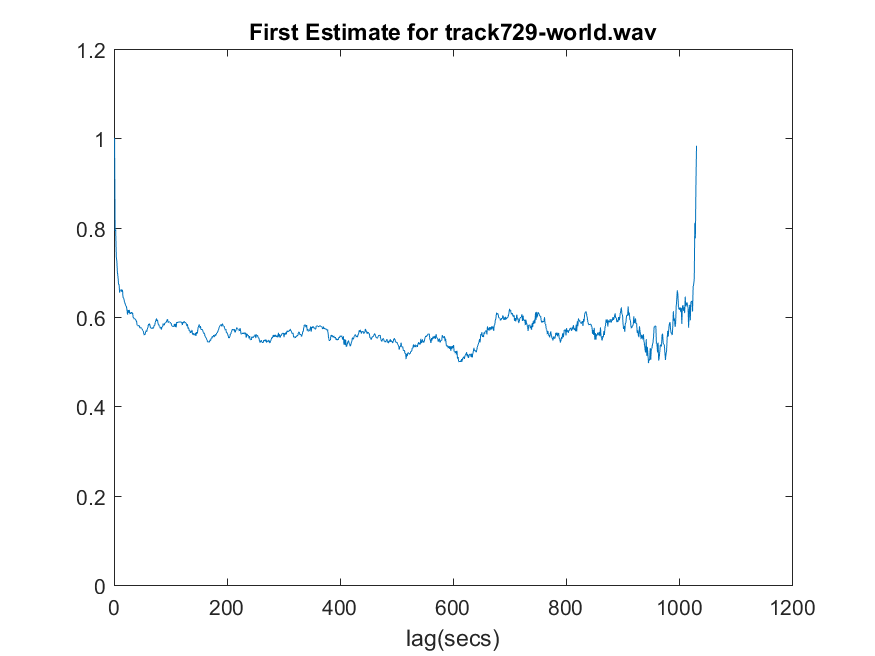


Fig 23: first estimate for track 707 Fig 24: first estimate for track 729

1. A better estimate of the rhythm

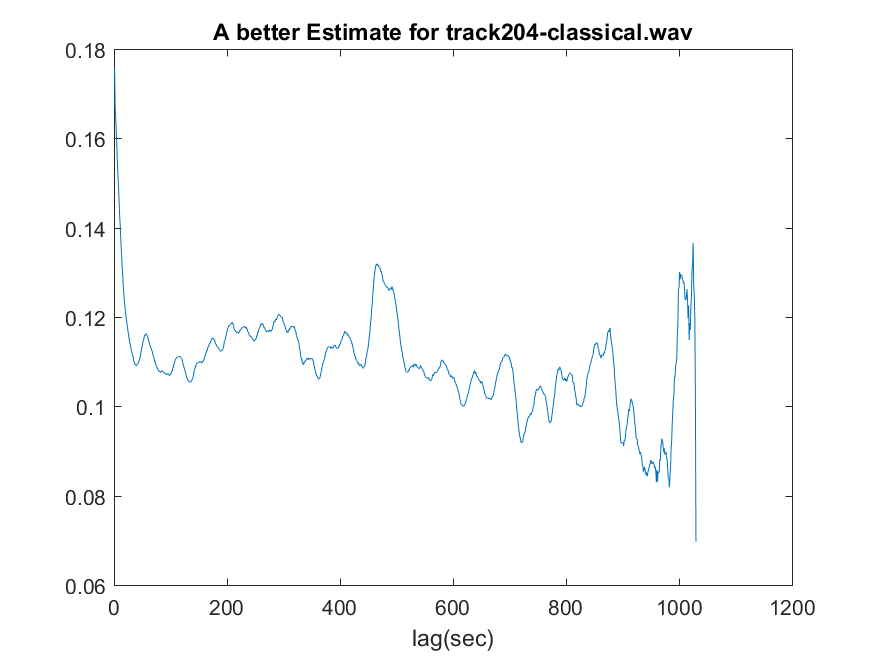
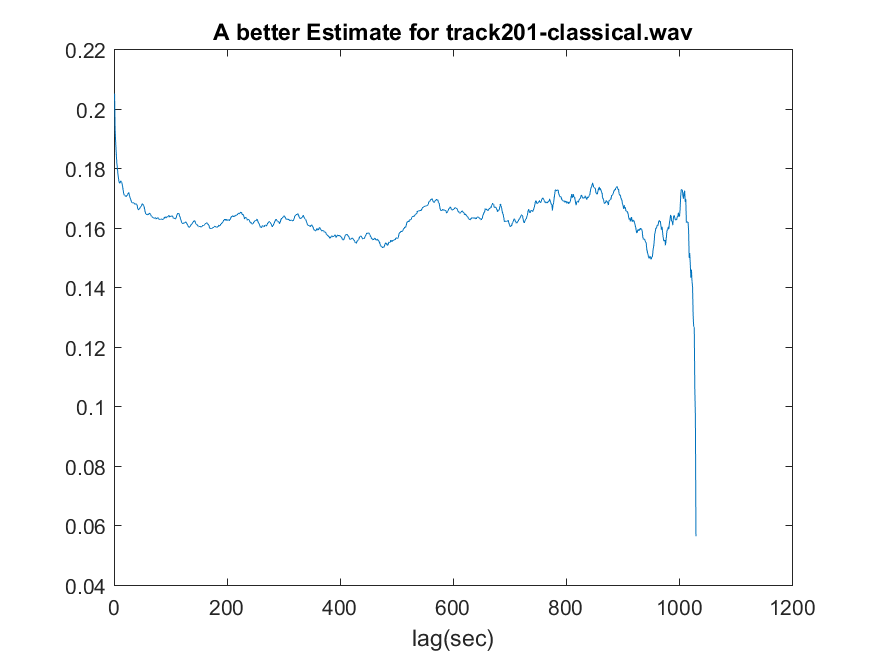


Fig 25: better estimate for track 201 Fig 26: better estimate for track 204

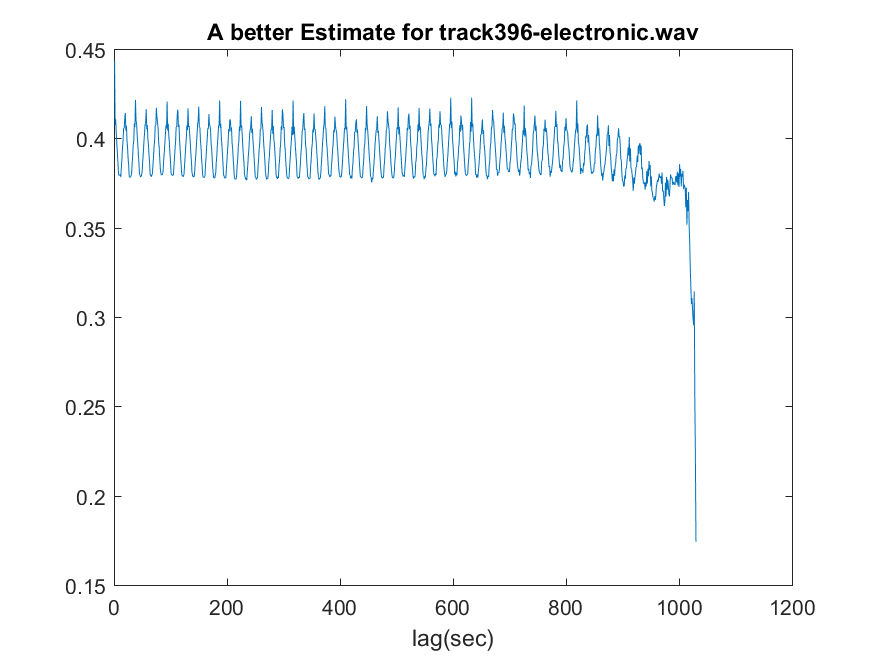
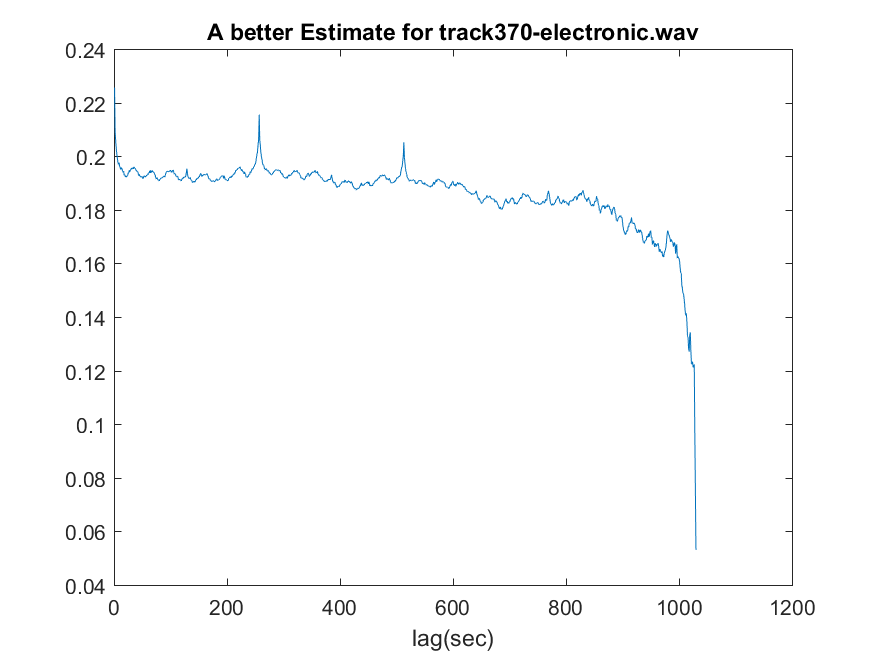


Fig 27: better estimate for track 370 Fig 28: better estimate for track 396

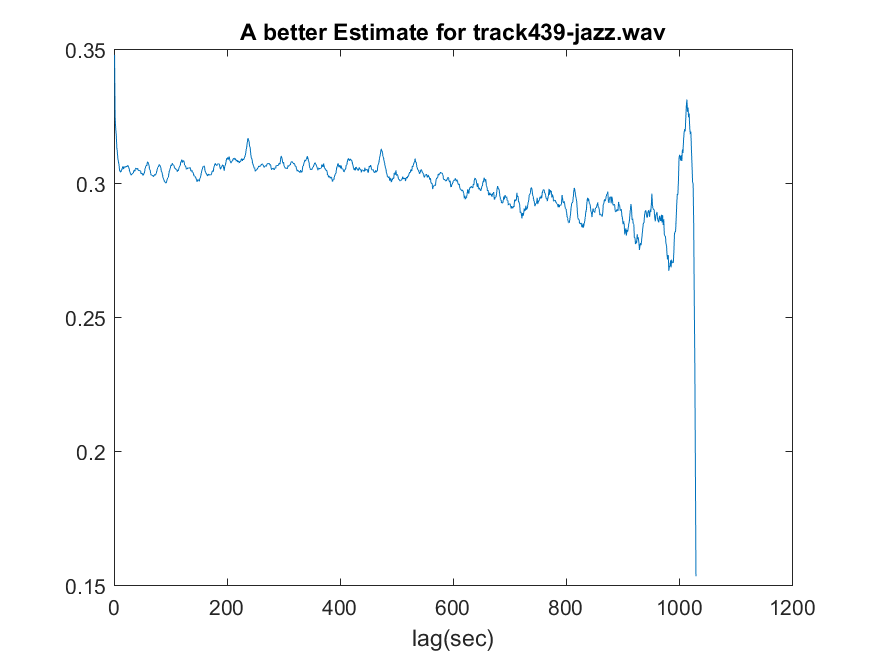
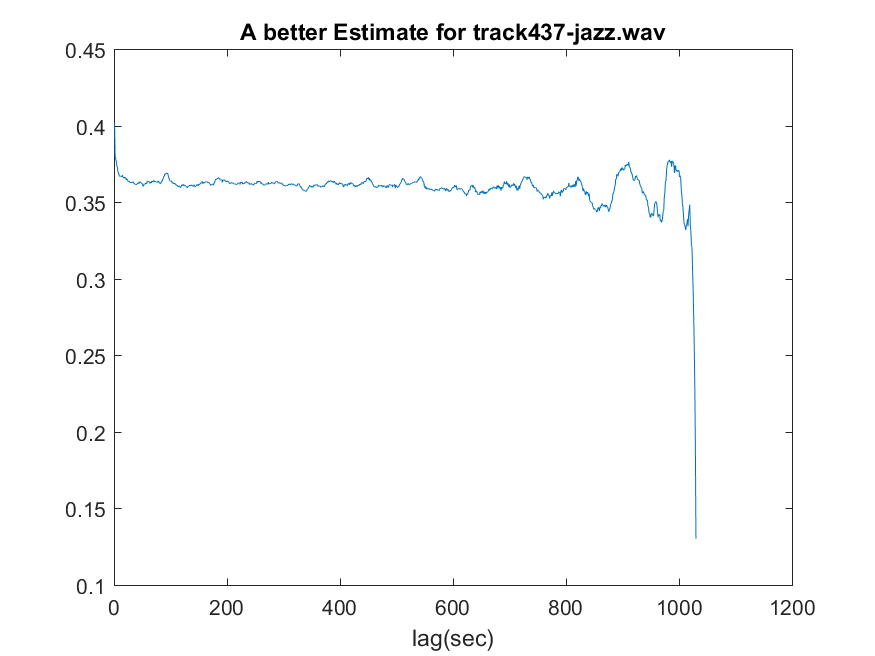


Fig 29: better estimate for track 437 Fig 30: better estimate for track 439

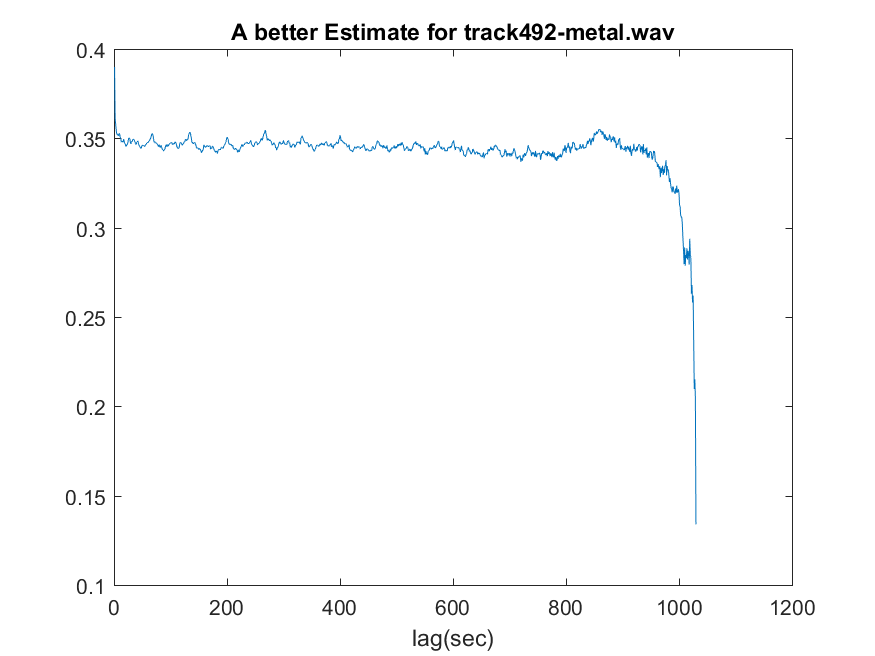
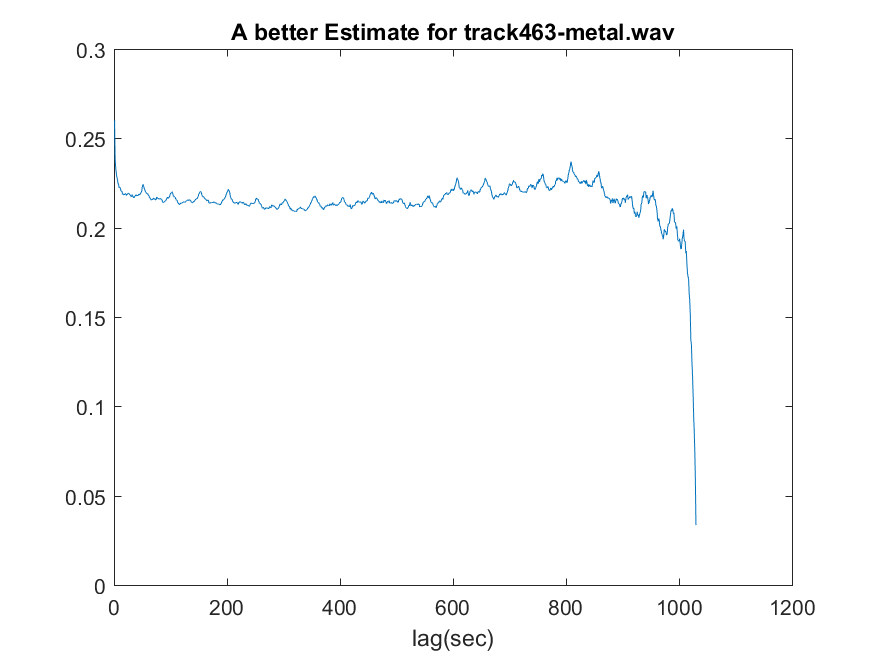


Fig 31: better estimate for track 463 Fig 32: better estimate for track 492

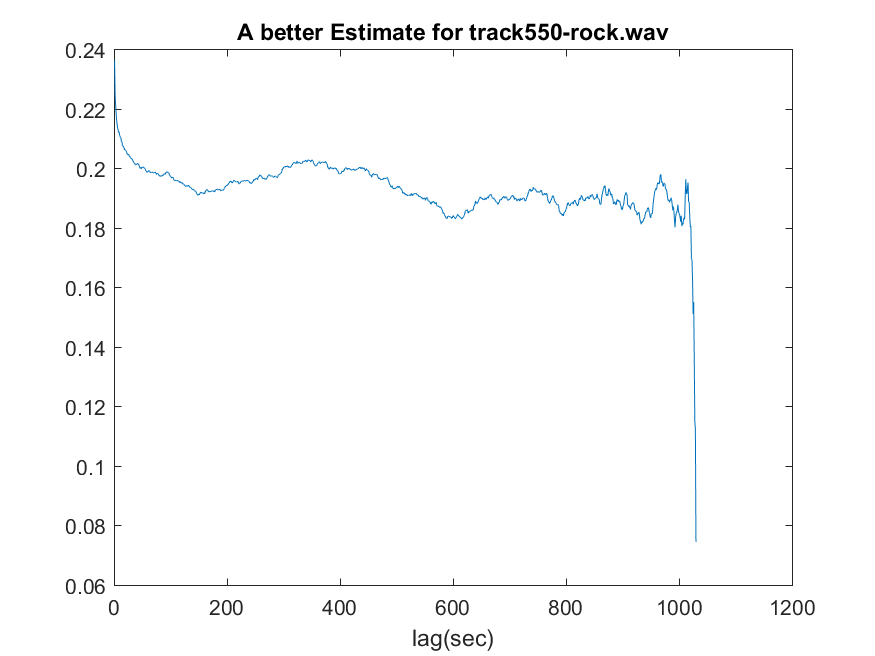
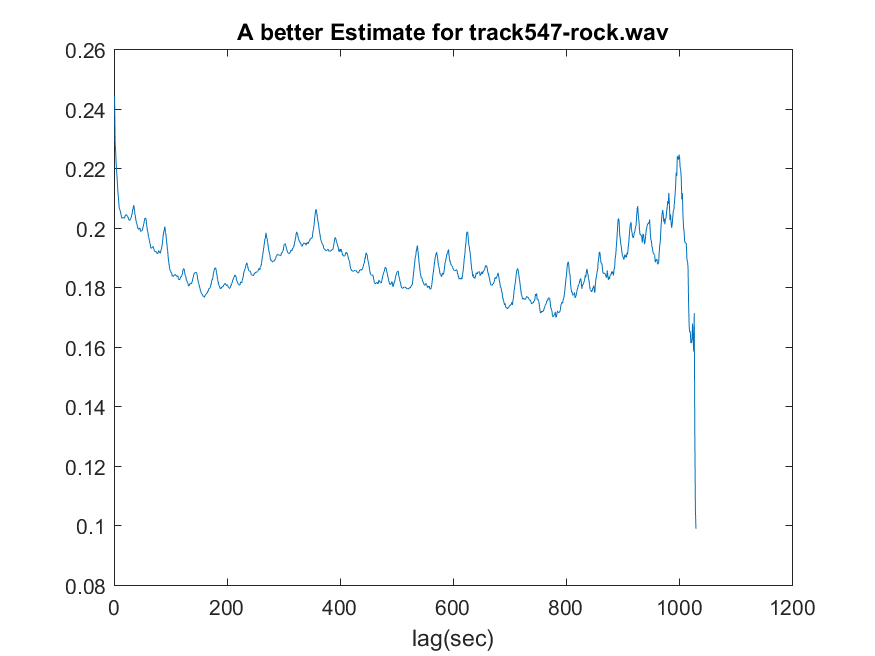


Fig 33: better estimate for track 547 Fig 34: better estimate for track 550

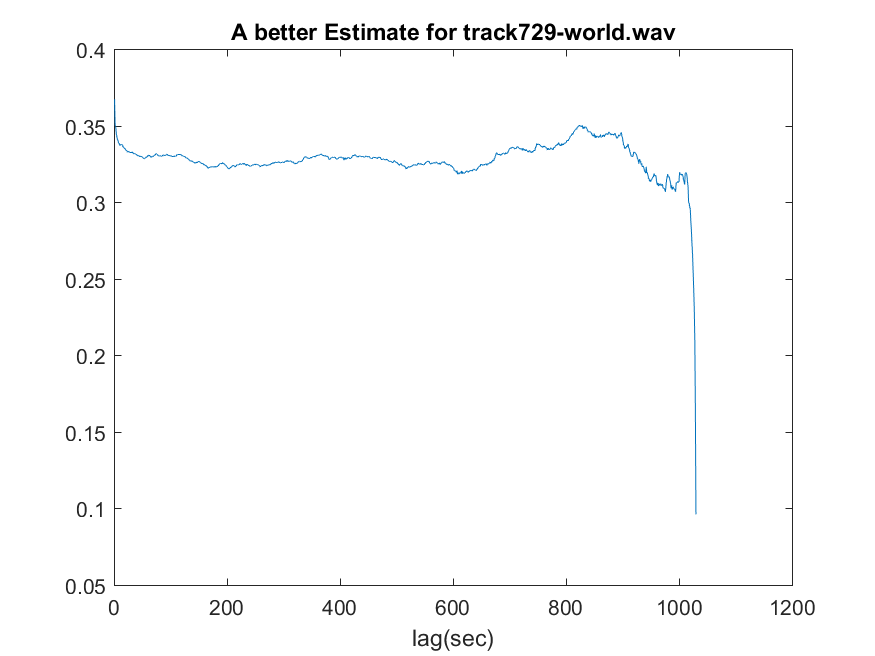
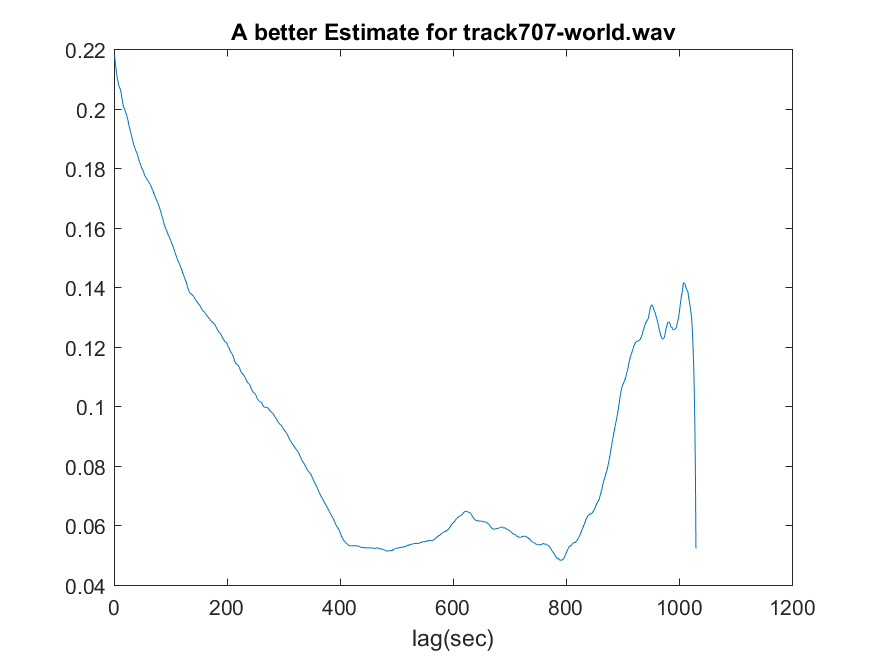
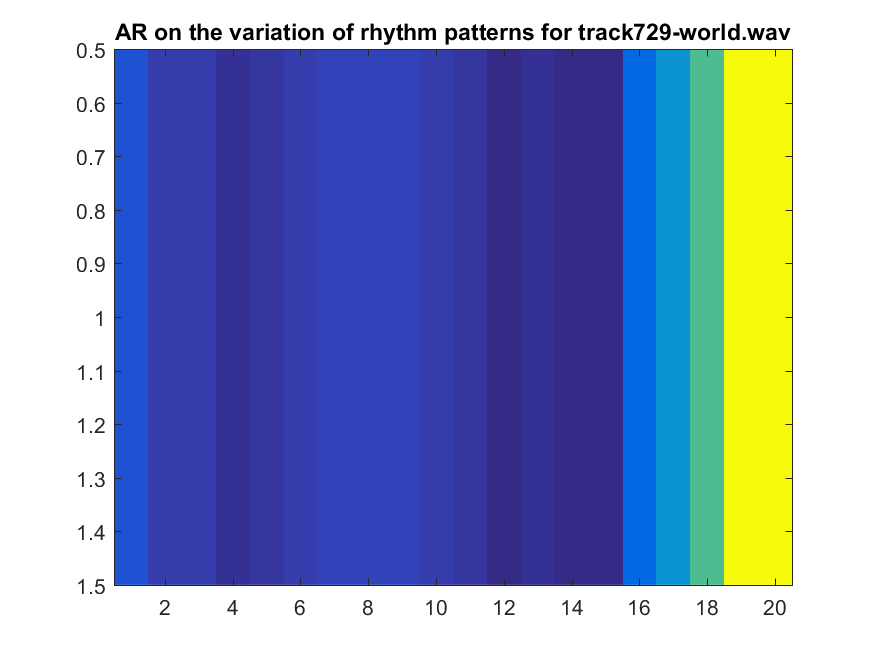
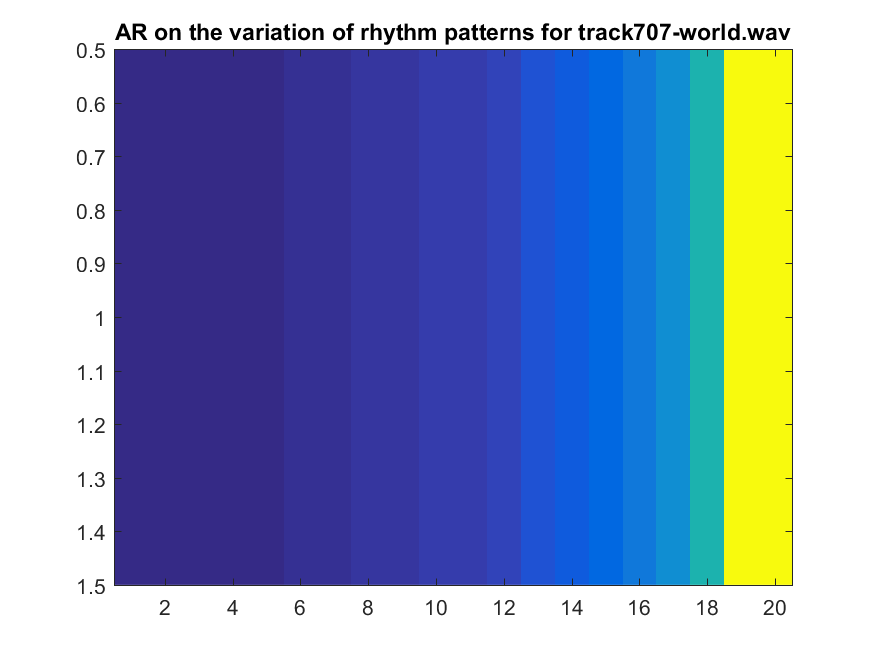
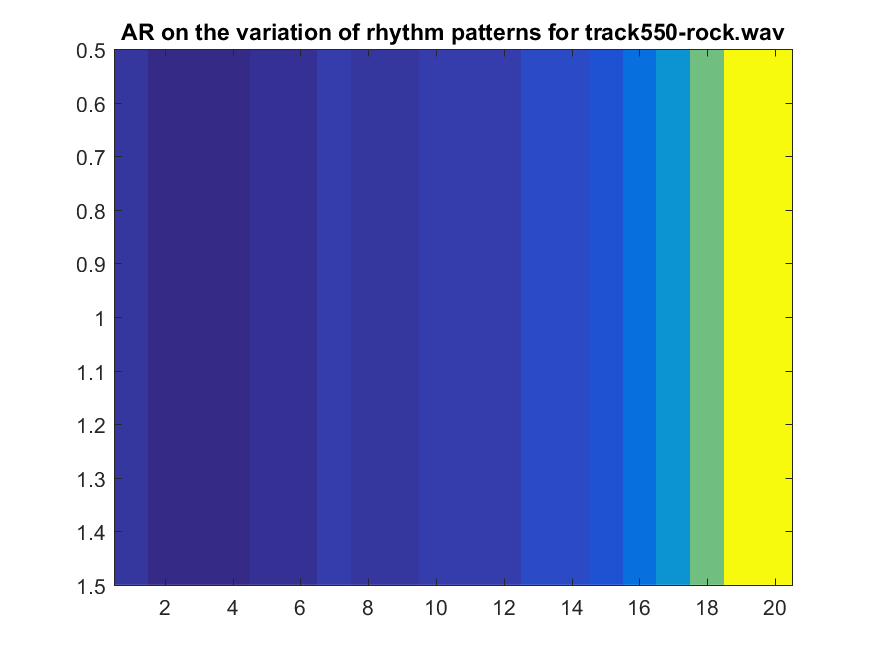
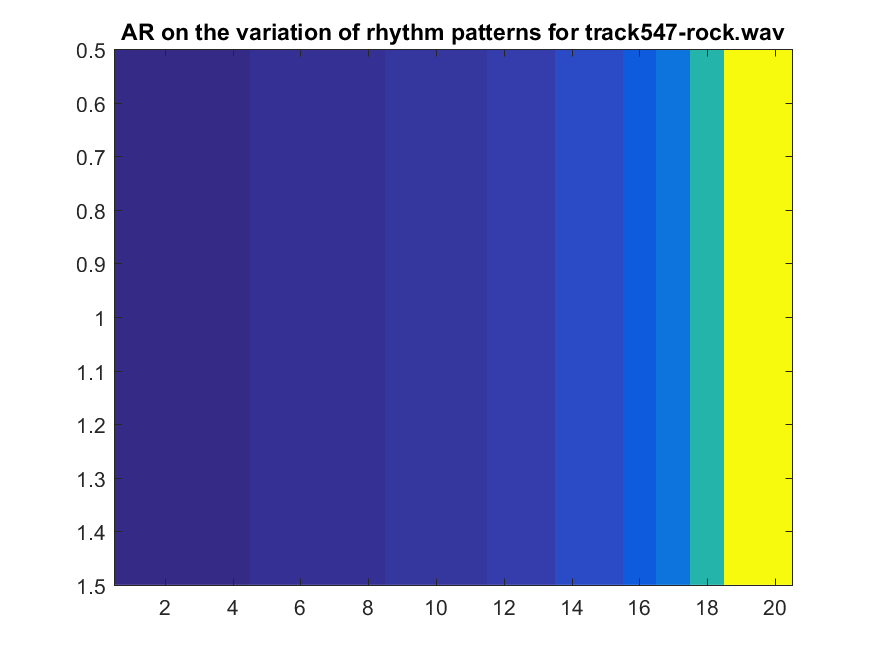
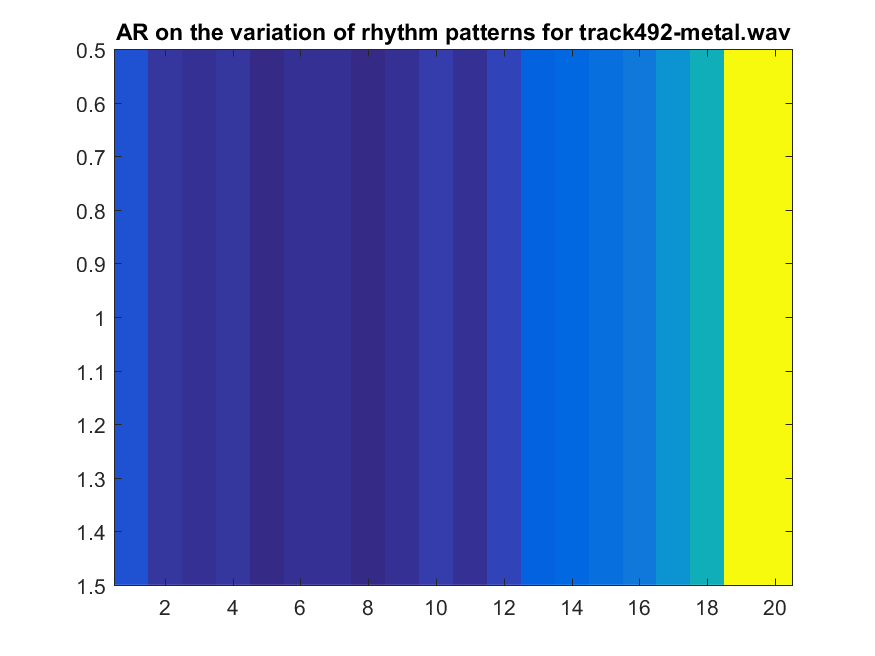
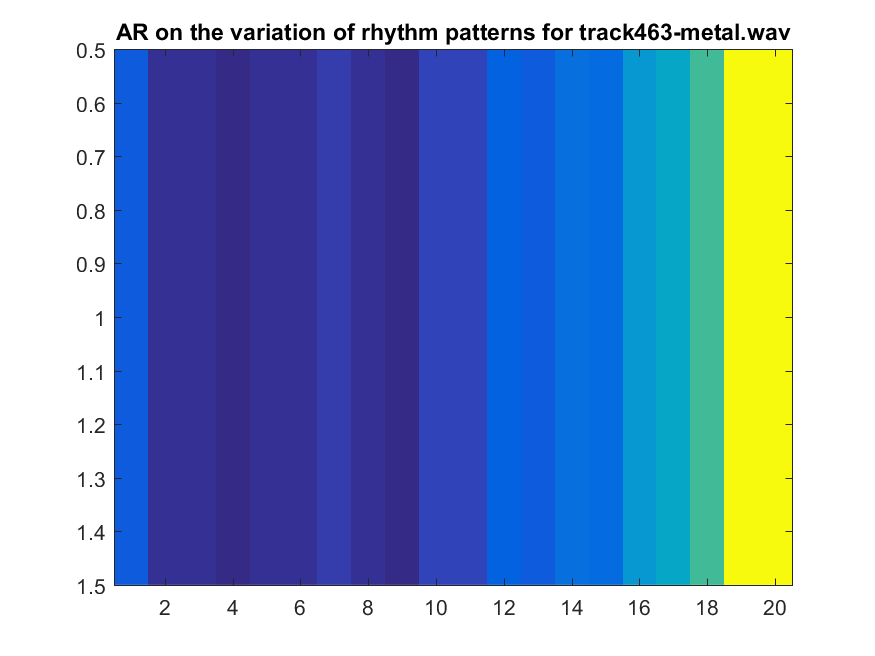
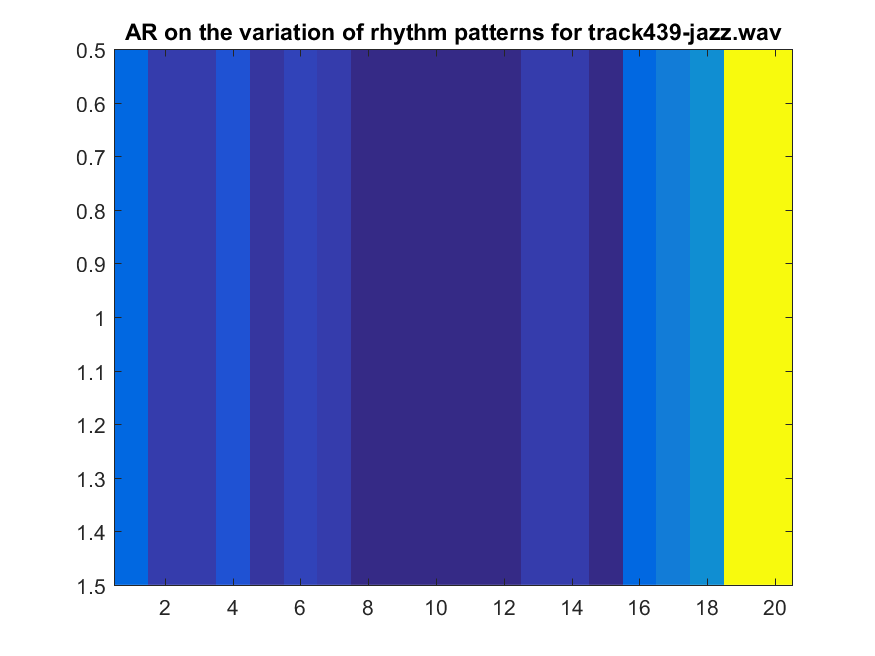
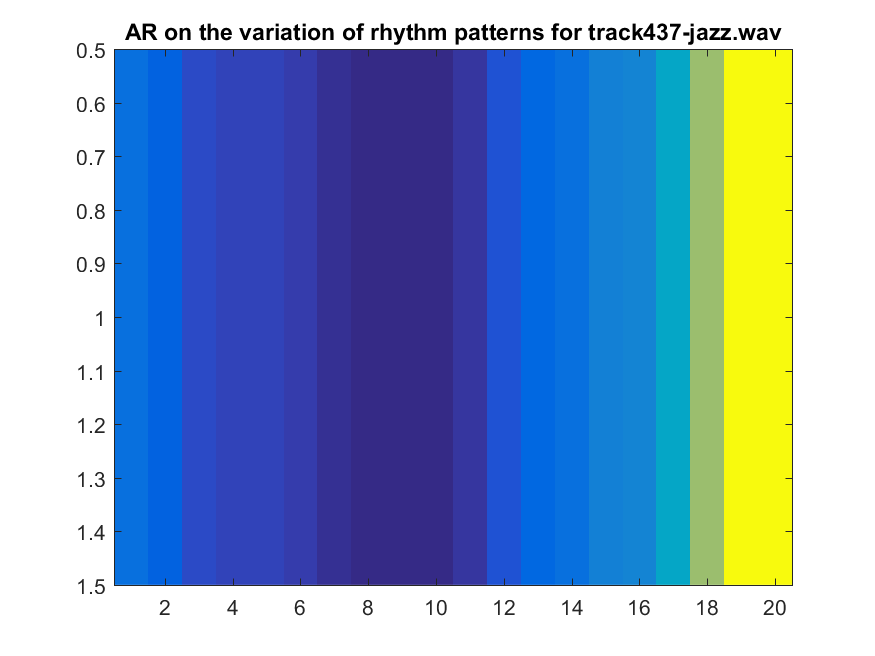
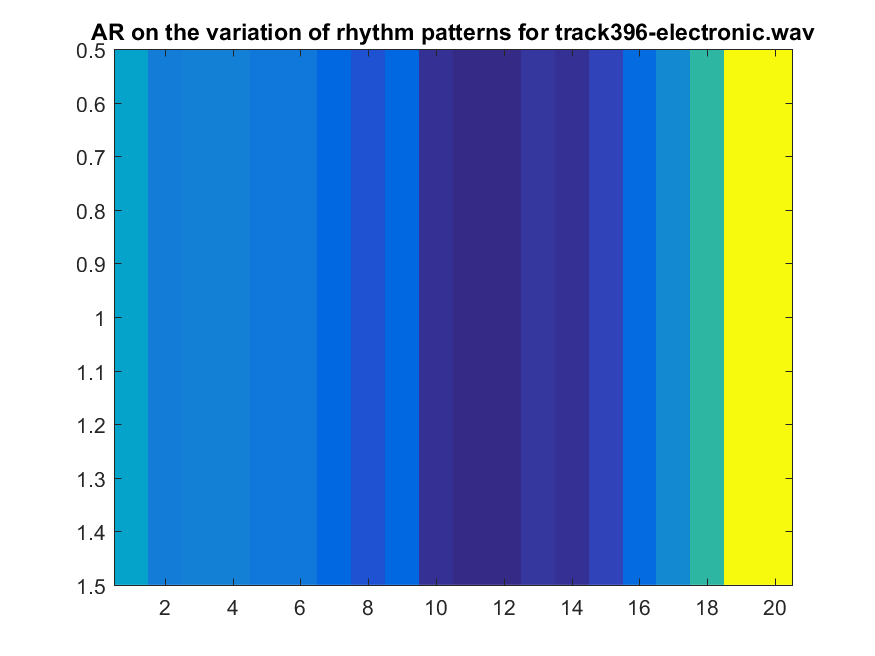
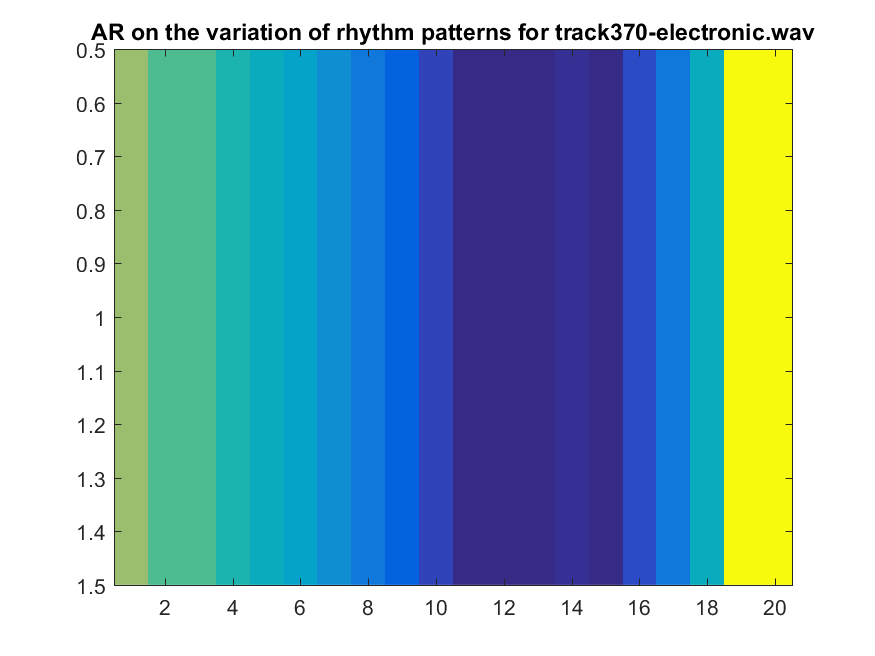
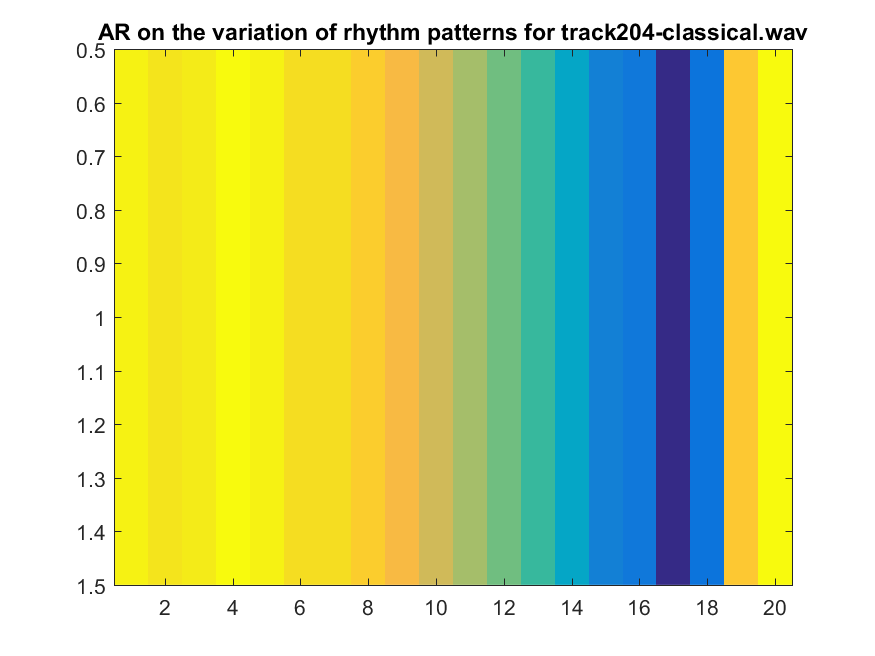
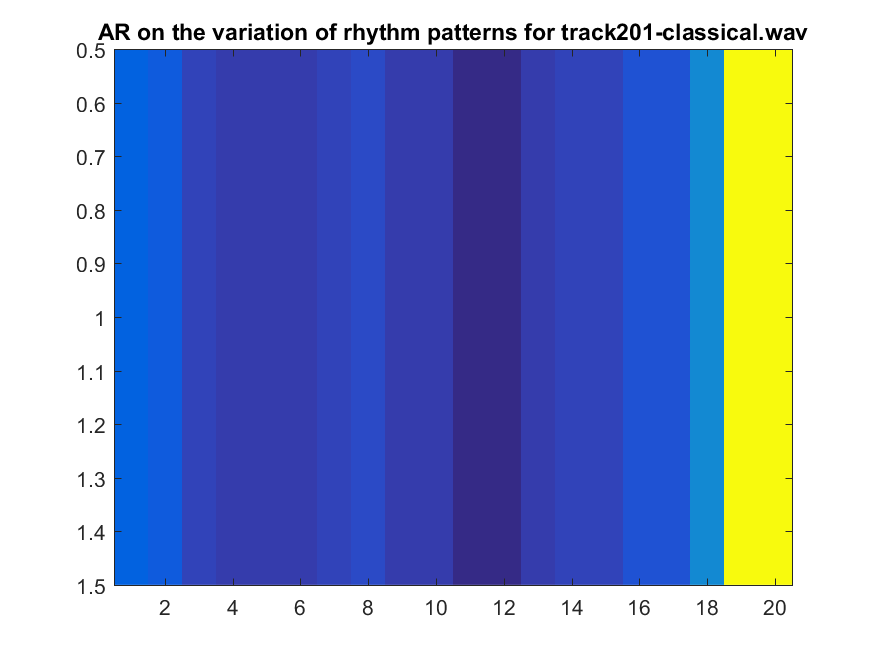


Fig 35: better estimate for track 707 Fig 36: better estimate for track 729

Computation of the rhythm index



Normalized Pitch Class Profile