20221024 Progress Report

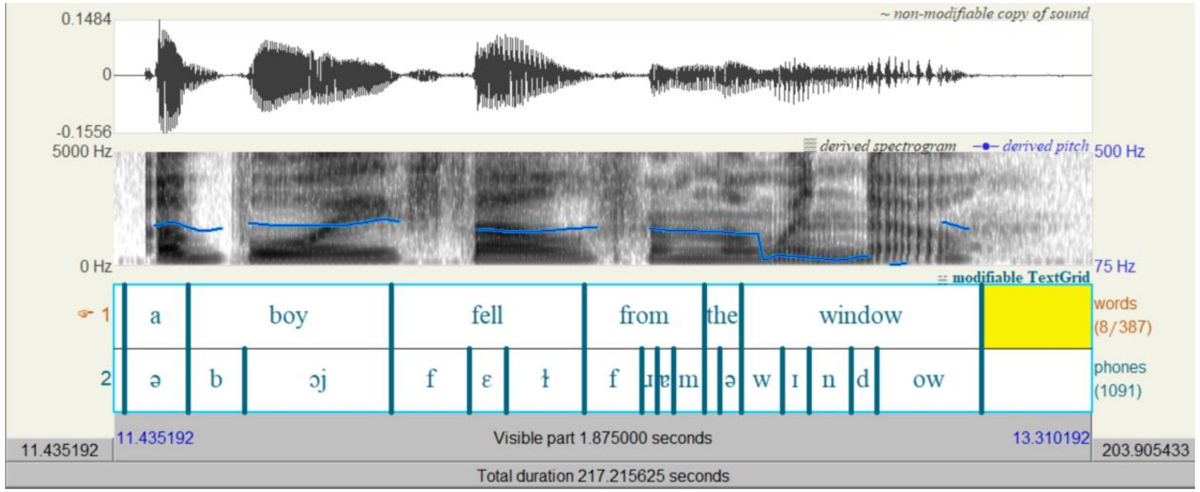
PingYao Feng

New data from real word speech!

Current work

Apply a record of American English containing several sentences to the MFA aligner, using english_mfa acoustic and english_mfa dictionary, and align the record into words and phones.

The next step is to segment .wav files according to the phone tier and get some topological features from these phones. To be more detailed, I will use some common vowels and consonants(shown left) in English, crop .wav files for each phone, and get a series of persistence diagrams for them.




```
vowel_phones=['ɔj','ɛ','ə','ɪ','aj','ɑ','æ','i','o','ʊ','aw','e','u','a']
consonant_phones=['b','f','m','ɹ','ð','w','h','p','t','z','n','g','dʒ','s','ʃ','v','l','ŋ','k','θ','j','tʃ','ʒ','d']
```

Questions

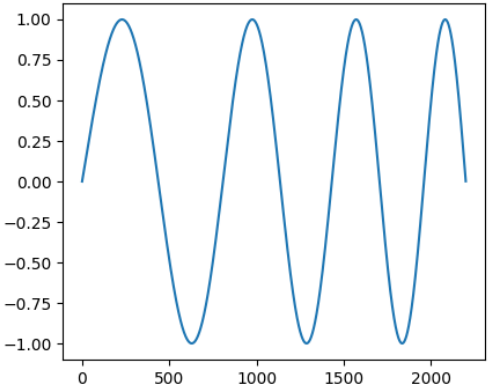
1. Are there any differences between the classification of vowels and consonants in linguistics and in the industry? If there are, why do those differences come into being?
2. What makes you a headache while dealing with speech data? There might be some drawbacks in current speech analysis. These may give TDA(topological data analysis) a try.

Questions continue

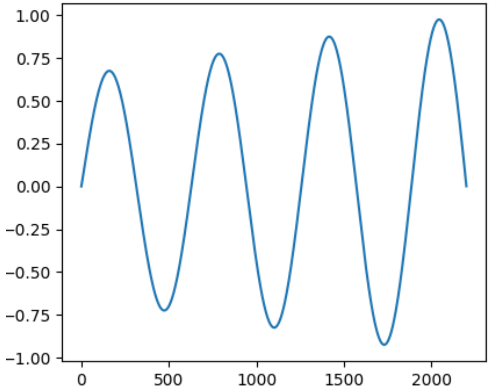
3. Normally how much and how long it takes to train a language model(English acoustic and dictionary for example) for industrial propose? Is there any way to make it cheaper and faster?

Progress Report

While the frame size is small enough, the changes in the wave of audio are one of the following:



1. Frequency

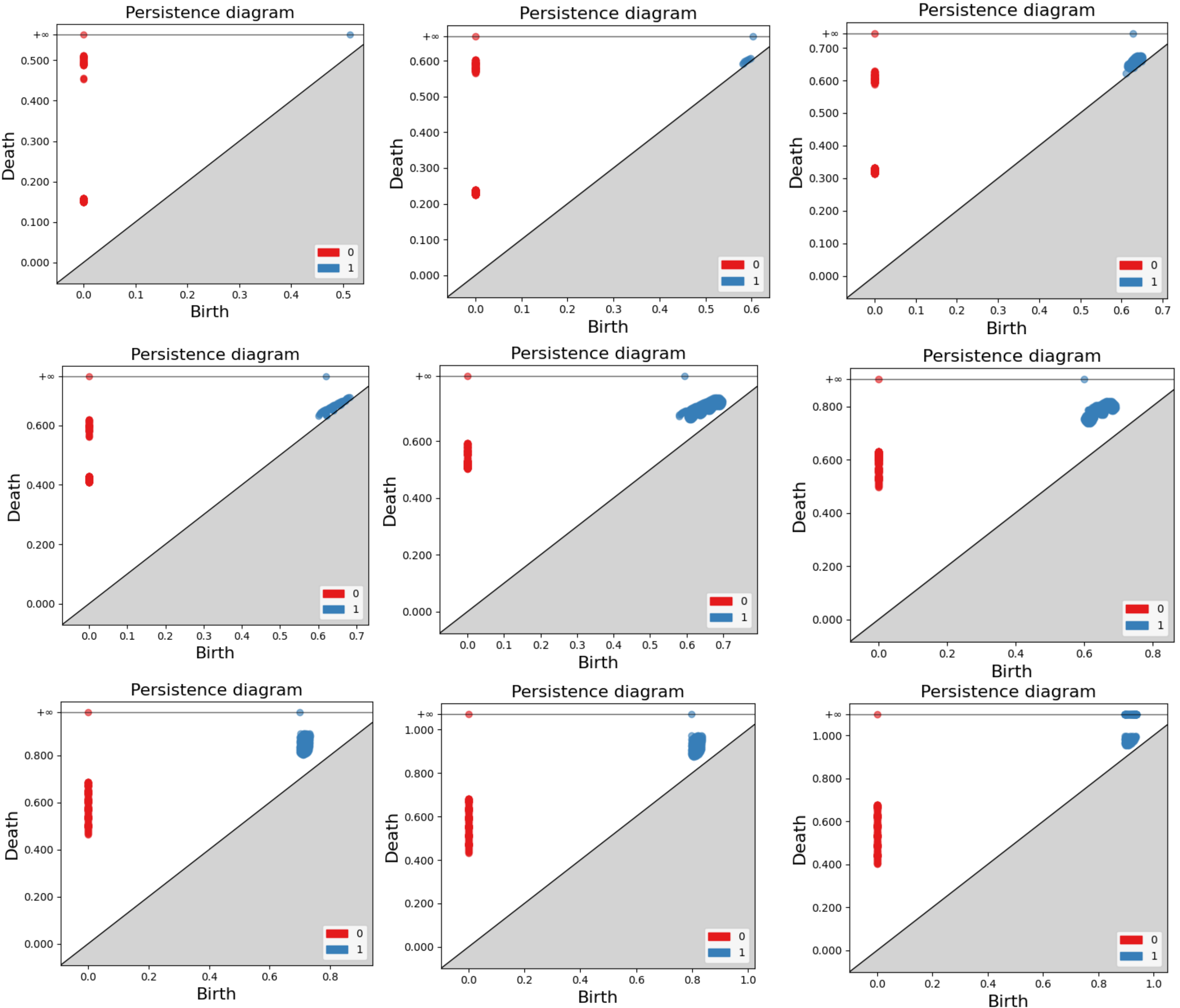


2. Amplitude

Persistent diagram can encode these small changes, and make this into a diagram. By observing the changes in the diagram, we can somehow learn how the wave of audio is changing without listening to the audio.

Amplitude

how persistent diagram change with amplitude?

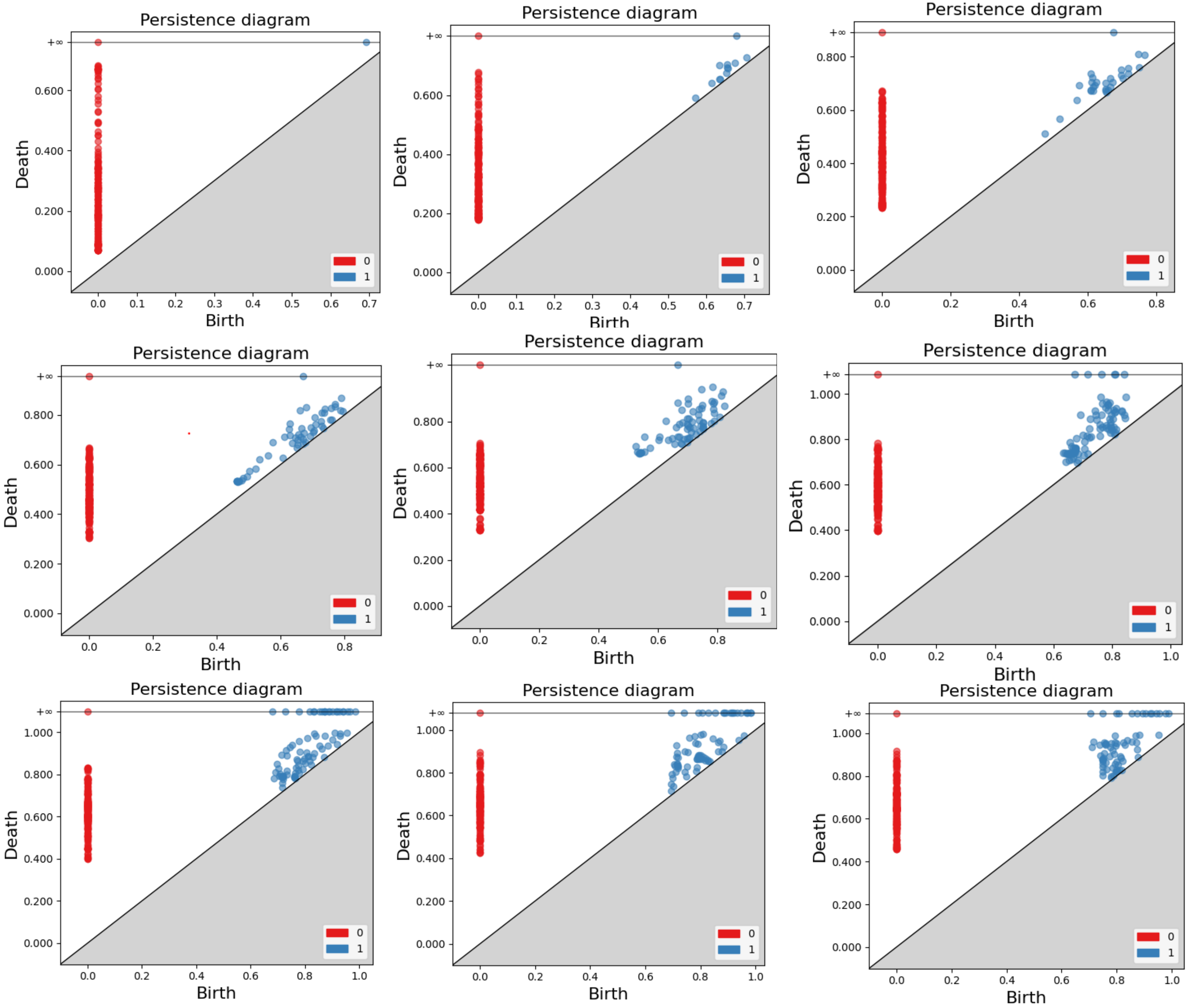



```
a=np.arange(0,7*np.pi,0.01)
for i in range(10):
    t=np.linspace(0.5+i/20,1,len(a))
    data_case1=t*np.sin(a)
```

How radical the amplitude is changing can be encoded into 0 and 1-dimensional persistent diagrams(corresponding to the red point and blue point, respectively).

Frequency

how persistent diagram change with frequency?




```
a=np.arange(0,7*np.pi,0.01)
for i in range(10):
    t=np.linspace(0.7+0.3*i/10,1,len(a))
    data_case1=np.sin(t*a)
```

How fast frequency is changing can also be shown in the persistent diagram. Note that it’s really different from the case in amplitude. Also, note that persistent diagram endures large frame size and rapid changes in both frequency and amplitude.