Who's the Bird behind the chirp?

Kevin Zecchini

02/19/2016

Motivation

Can we classify sounds by their acoustic characteristics?

Why birds?

- Non-invasive way to track wildlife patterns
- We hear birds every day
- Can lead to classification of more complex soundscapes

Data:

- Collected from Xeno-Canto
- Only ‘A’ rated songs used
- Six species to classify
- 240 total recordings

Dataset

Barred Antsrike

American Robin

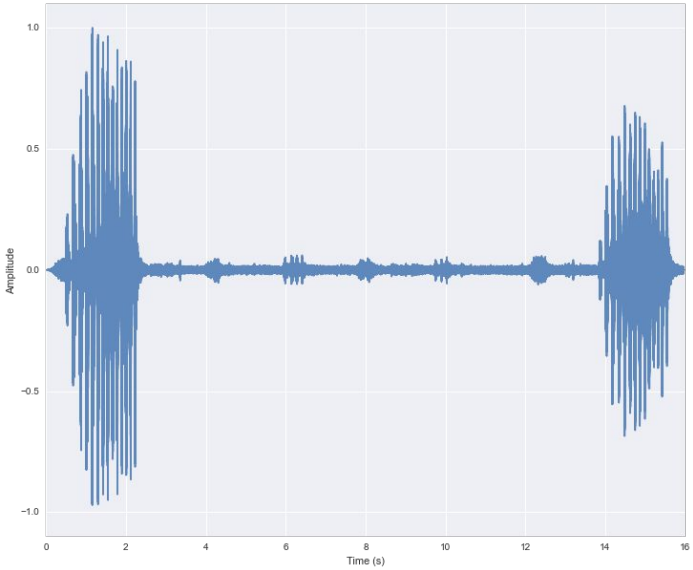
Northern Cardinal

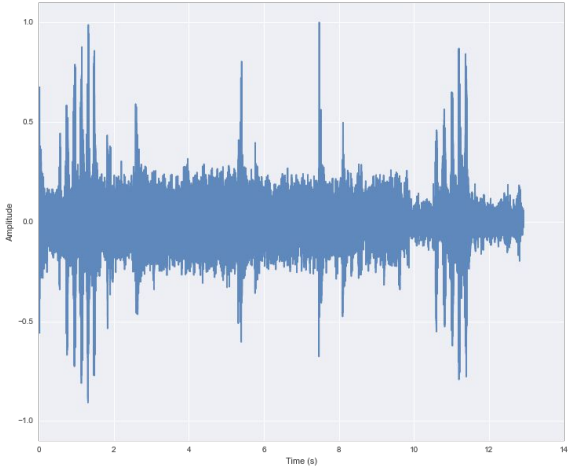
Great Antshrike

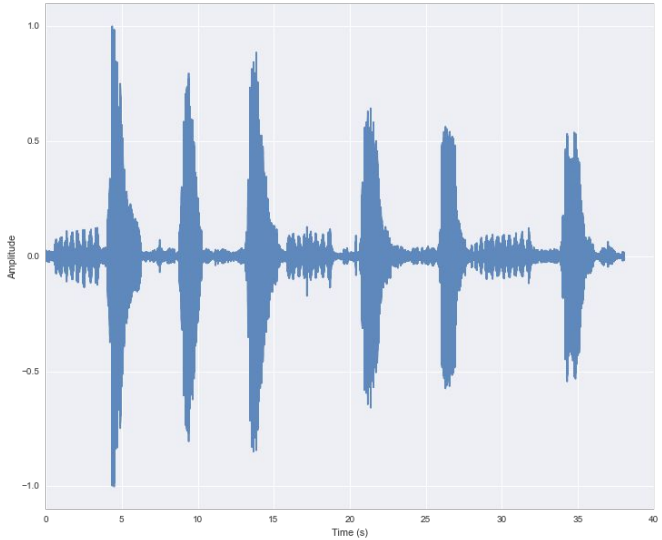
Dusky Antbird

House Finch

Signal Processing

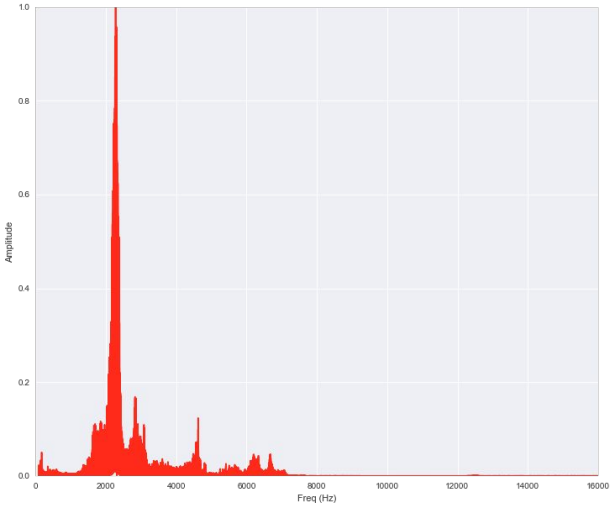


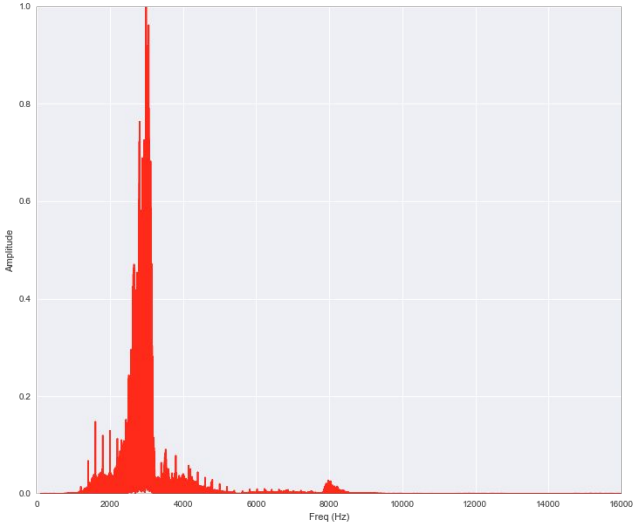


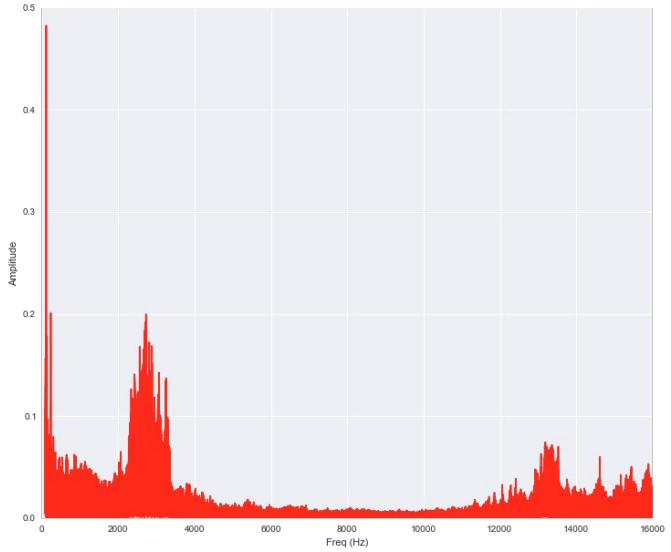


This is all the same bird - the dusky antbird!

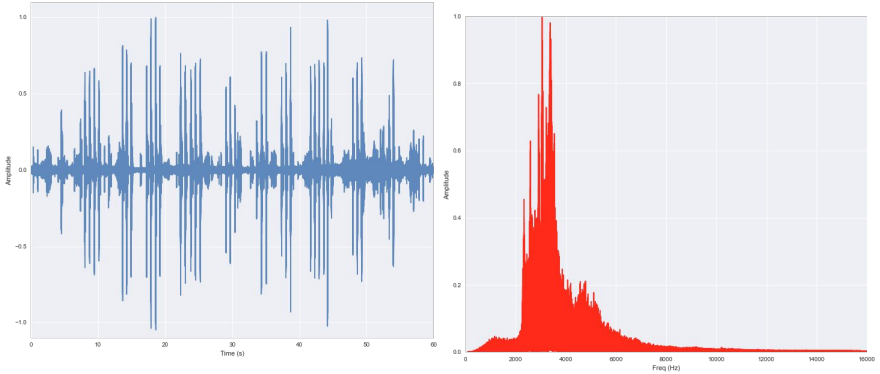
Signal Processing



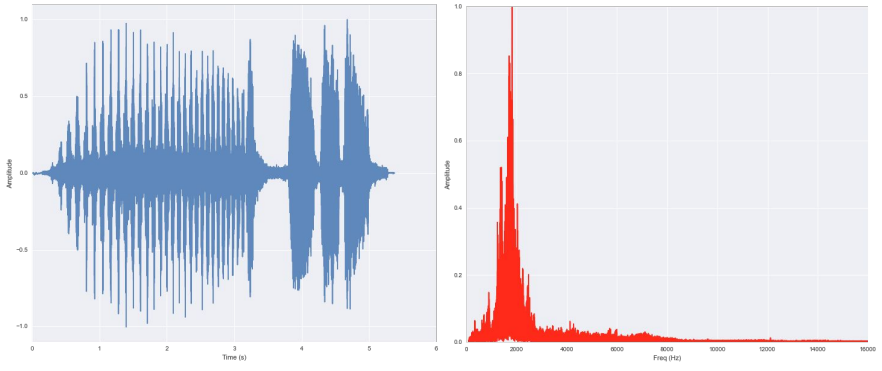




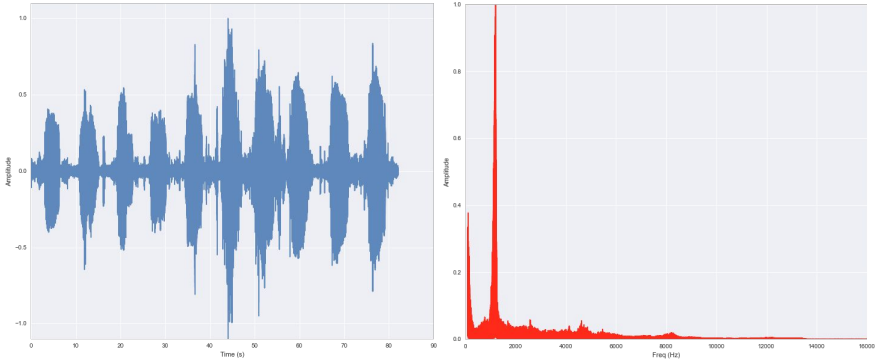
Northern Cardinal



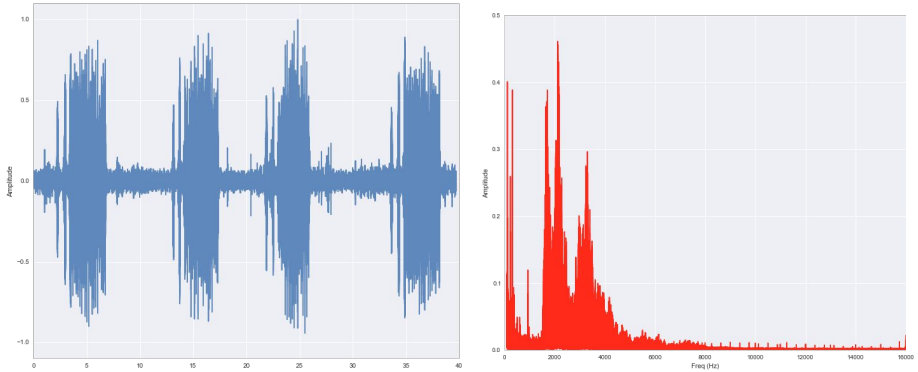
Barred Antsrike



Great Antsrike



American Robin



Extract acoustic qualities - tone, timbre, brightness/color, tempo, rhythm

Starting Features

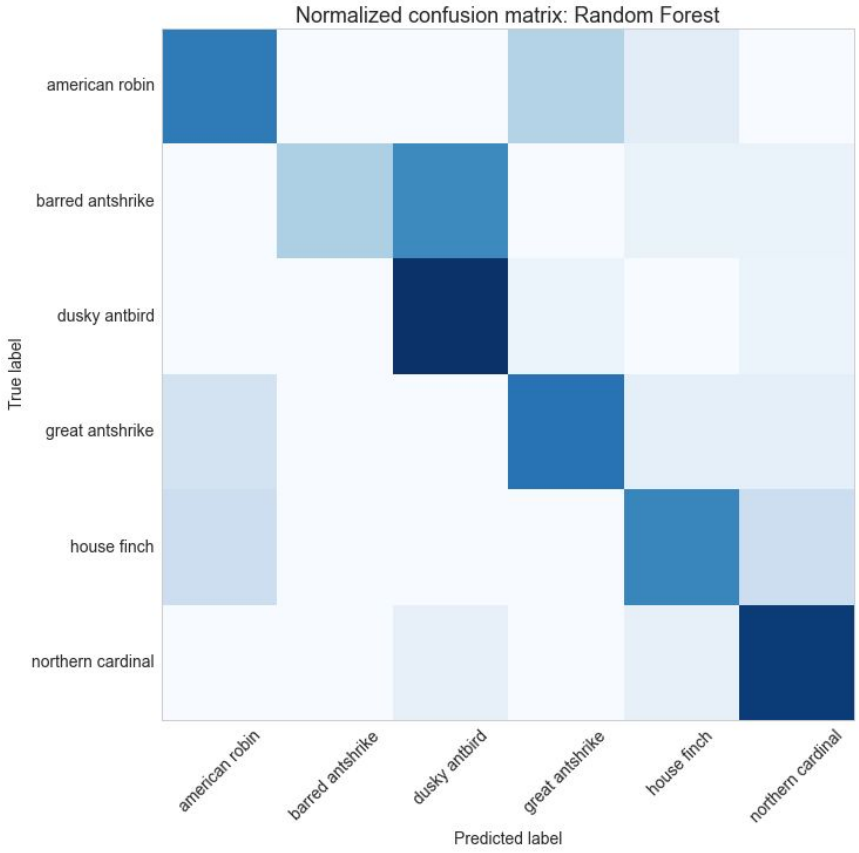
- Entire Signal:
  - 15 frequency bins of avg. fourier amplitude values
  - 4 bins of amplitude distribution
  - Freq. of max amplitude
- Aggregate over 25ms frames:
  - Root mean square energy
  - Avg. number of zero crossings
  - Std. of Spectral Centroid
  - 13 avg. Mel Frequency Cepstral Coefficients

After Feature Selection

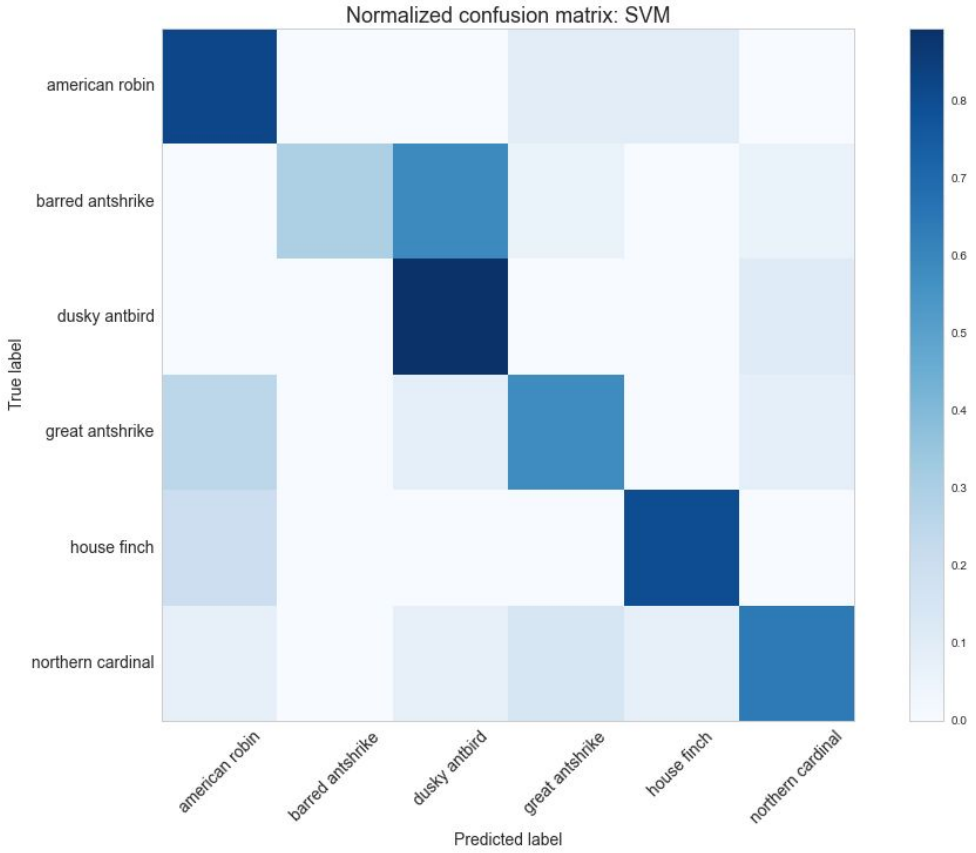
- 5 frequency bins of avg. fourier amplitude values
- Freq. of max amplitude
- 3 bins of amplitude distribution
- First and thirteenth avg. MFCC
- Avg. number of zero crossings

Chirp! Chirp! - Who's There?

Accuracy = 0.692



Accuracy = 0.654

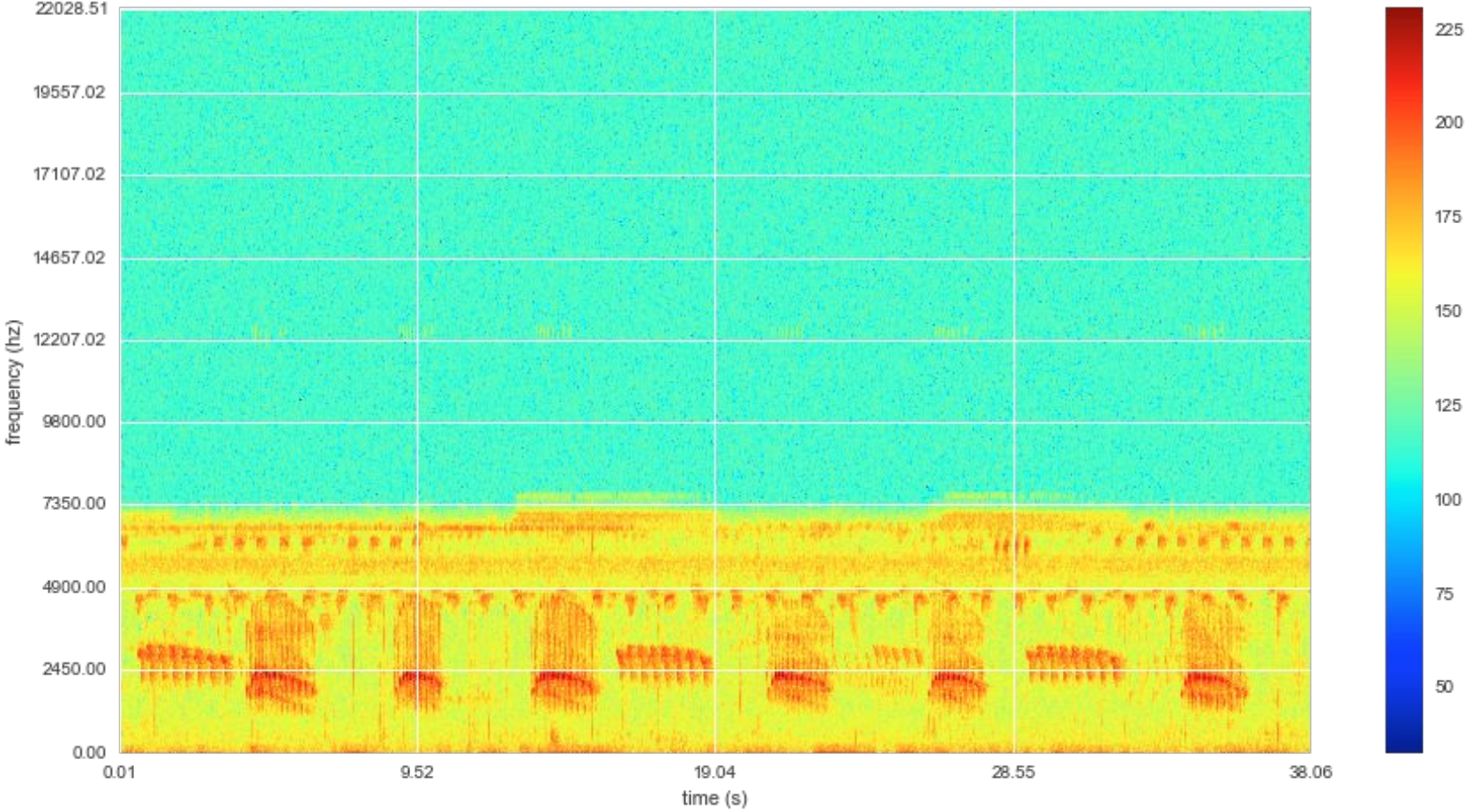


Challenges/Next Steps

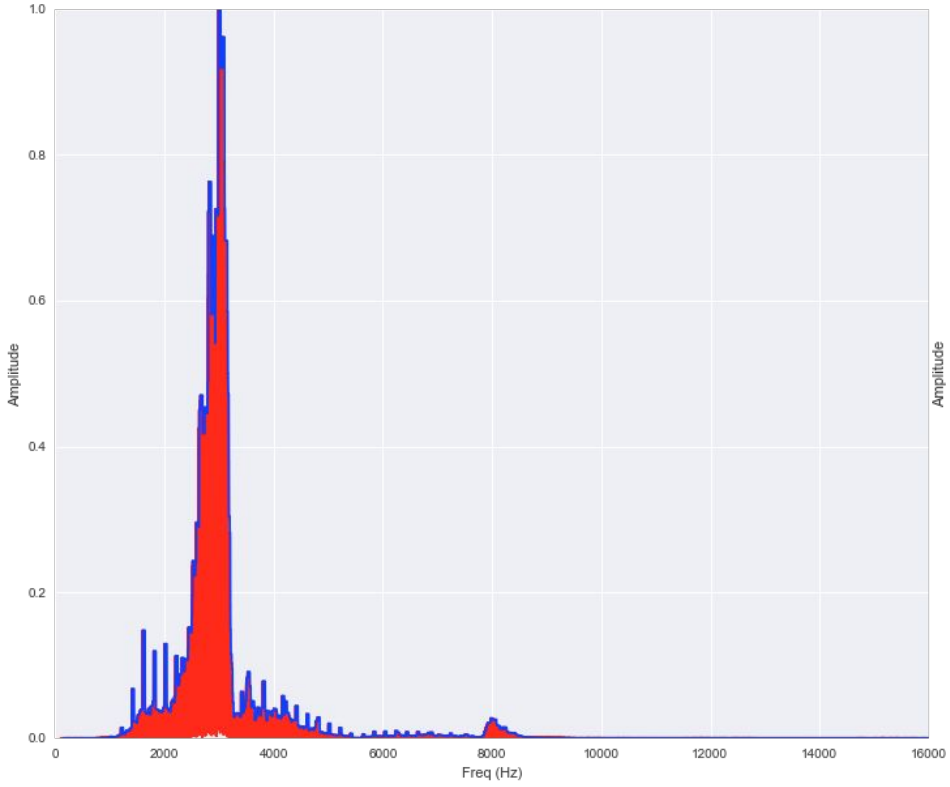
- Major challenge: the time domain
  - Dynamic Time Warp and Auto Correlation
  - Align on the correct part of the song
  - Songs consist of permutation of notes/sequences
  - Account for tempo, rhythm
- Explore feature selection/creation in more depth
  - Distinguishing birds with similar frequency range
  - Unsupervised learning
- Use neural network or deep learning
- Add more birds!

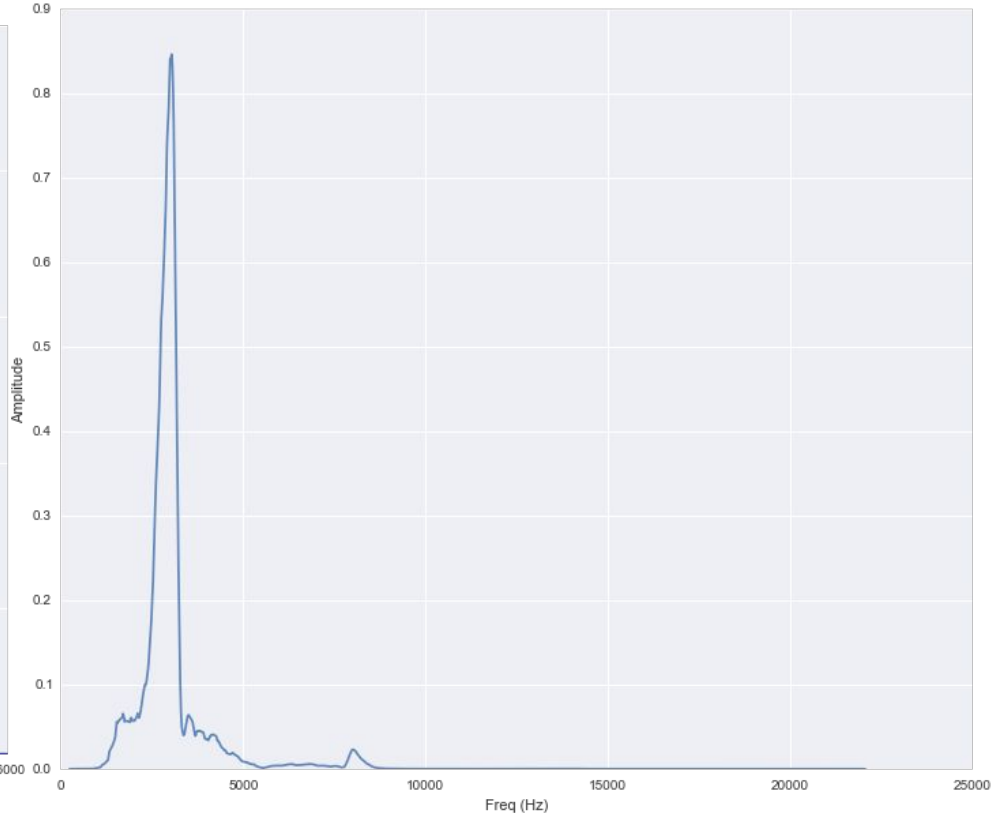
Questions?

Spectogram Space



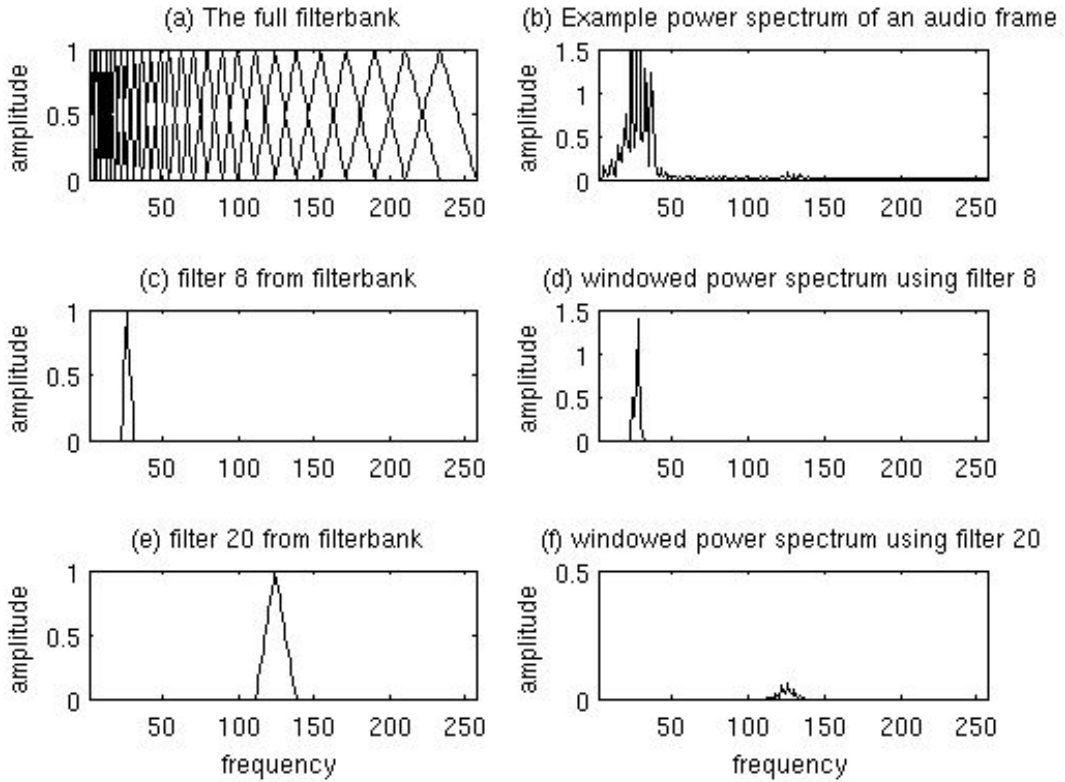
Signal Processing



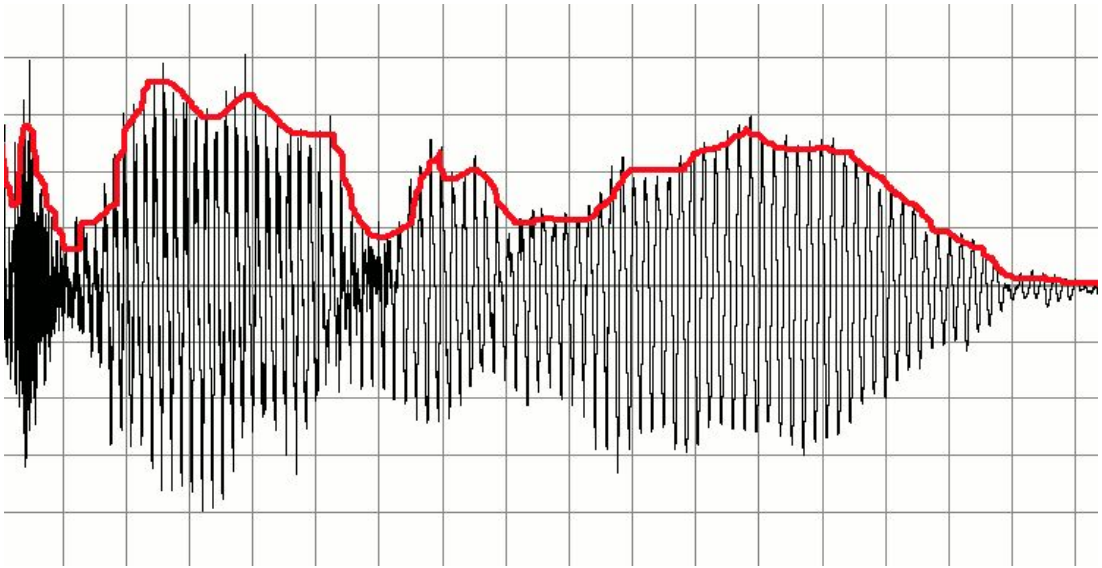


MFCC

$$M(f) = 1125 \ln(1 + f/700) \quad (1)$$



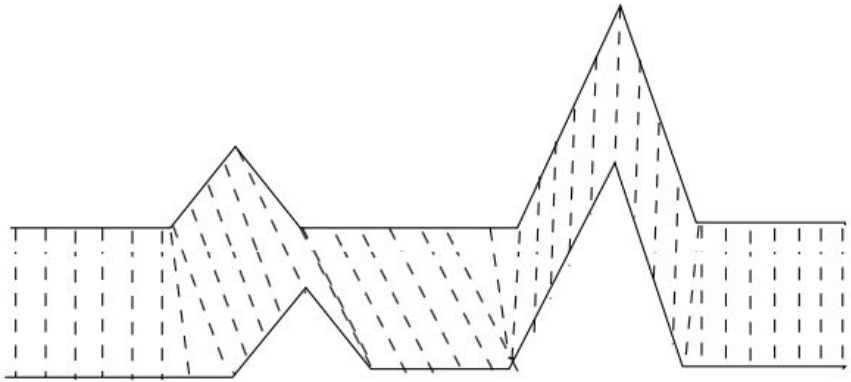
Timbre - wave envelope



Dynamic Time Warp/Auto correlation

DTW = Finding the best path between two time series

- Sequences are ‘warped’ non-linearly in time

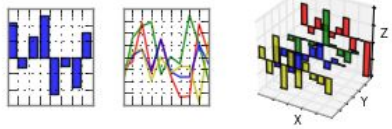


Contents

Requests

pandas

$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$



SM StatsModels

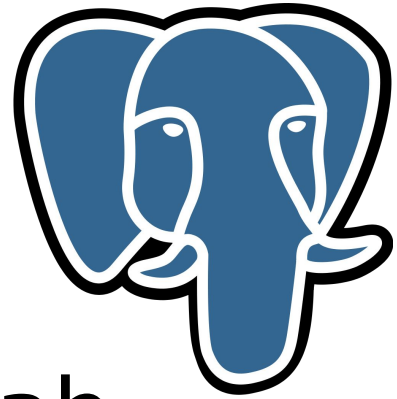
Statistics in Python

NumPy

matplotlib

mlpy

machine learning py



scikit learn

Librosa

Neurolab

Resources

https://www.researchgate.net/publication/224266043_Feature_set_comparison_for_automatic_bird_species_identification

http://www.academia.edu/2272643/Classification_of_Birds_using_FFT_and_Artificial_Neural_Networks

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4106198/

http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/

http://web.engr.oregonstate.edu/~xfern/bird-icdm09.pdf