Topic

Spectrogram
Chromagram
Cesptrogram
Short time Fourier Transform

- Break signal into windows
- Calculate DFT of each window
The Spectrogram

- A series of short term DFTs
- Typically just displays the magnitudes of $X$ from 0 Hz to Nyquist rate

spectrogram(y,1024,512,1024,fs,'yaxis');
Equal Temperament

• Octave is a relationship by power of 2.
• There are 12 half-steps in an octave

\[
f = 2^{\frac{n}{12}} f_{ref}
\]

number of half-steps from the reference pitch

frequency of desired pitch

frequency of the reference pitch

Spiral Pitch representation
Chroma: Many to one

- Chroma = \log_2(\text{freq}) - \text{floor} (\log_2(\text{freq}))
- Chroma periodic in range 0 to (almost) 1
- Chroma map on to pitch classes
Making a Chromagram

• Decide how to quantize (bin) the chroma range.
  – 12 pitch classes? 120 bins? Equal temperment?
• Make a spectrogram
• For each time-step in the spectrogram
  – find the chroma for each frequency from 0 to N/2
  – Sum the amplitude of all frequencies with the same chroma bin
    • (Some chromagrams also add in the energy from the odd harmonics)
  – Place that value in the chroma bin
Overtone Series

- Approximate notated pitch for the harmonics (overtones) of a frequency
A fancier chromagram

- For complex sounds (like the bassoon example from class) you might want to consider adding up energy from more harmonics than just the octaves (1f, 2f, 4f…etc).
- Try taking the energy from the 3rd, 5th and 7th harmonics as well.
Chromagram of Clarinet
Chromagram of Clarinet
Mel Scale

- Stevens, Volkmann and Newmann (1937)
- A scale of pitches judged by listeners to be equidistant.
- The reference point:
  - 1000 mels = 1000 Hz at 40 dB SPL
- Below 500Hz mel ~= hertz
- Above 1000 Hz mel ~= log(hertz)

Mel Filter Bank

- Filters spaced equally in the log of the frequency.
- Mels are (more or less) related to frequency by…

\[
f_{mel} = 2595 \log_{10} \left( \frac{f}{700} + 1 \right)
\]

- Edge of each filter = center frequency of adjacent filter
- Typically, 40 filters are used

Source-Filter Model

\[ x(t) \ast h(t) = y(t) \]

Convolution
The Cepstrum

• Filtering is
  – Convolution in the time domain
  – A product in the frequency domain
• What if we want to make it an addition operation?

\[
Y[k] = X[k] \cdot H[k]
\]

\[
|Y[k]| = |X[k]| \cdot |H[k]|
\]

\[
\log(|Y[k]|) = \log(|X[k]|) + \log(|H[k]|)
\]
The Cepstrum

• Filtering is
  – Convolution in the time domain
  – A product in the frequency domain
• What if we want to make it an addition operation?
• They do this by defining the **cepstrum**.

\[ Cep_x(q) = Z^{-1}(\log |X(z)|) \]

A frequency representation

What is the Cepstrum for?

- Invented for finding echoes (aftershocks) in seismograph data.

- If something is useful for finding echoes, it is useful for finding impulse response functions

- …which makes it useful for finding filter coefficients.

Let’s look at an example…

Some terms

• Spectrum
• Spectrogram
• Frequency
• Filtering
• Cepstrum
• Cepstrogram
• Quefrecency
• Liftering
The Cepstrum

• Gives information about rate of change in the different quefrency bands.
• Popular representation for speech and music
• Distinguishing FILTER from the SIGNAL
  – Some quefrencies represent the filter (what instrument), others represent the signal (what pitch)
• For these applications, the spectrum is usually first transformed to Mel Frequency bands.
• Result: **Mel Frequency Cepstral Coefficients (MFCC)**
Making a Mel Freq Cepstrogram

Here DCT = Discrete Cosine Transform
Let’s have a look!

• (Go to bassoon/tuba demo)