Revisited: Fundamental Frequency Detection/Pitch Tracking

Bryce Irvin & Matt McMullen
Computational Music and Audio Analysis
Professor Lerch
November 14, 2021

A3. If the blockSize = 1024 for blocking, what is the exact time resolution of your pitch tracker? Can this be improved without changing the block-size? If yes, how? If no, why? (Use a sampling rate of 44100Hz for all calculations).

$$\text{Equation 7.18: } f_Q = \kappa \frac{f_s}{\mathcal{K}} = \kappa * 43.0664 \, Hz$$

Without changing either the sample rate (f_s) or blockSize ($\mathcal{K}$), the frequency resolution cannot be improved at no cost; however if need be, the resolution can be effectively increased with the use of certain workarounds. For instance, the spectrum can be interpolated or the blocked audio can be zero-padded to increase the "block length" and thus increase the frequency resolution. Neither of these solutions are without error, though. Aside from the obvious potential error from interpolation, the zero-padding reshapes the spectral leakage of the transformation.

E1. In a separate function executeassign3(), generate a test signal (sine wave, f = 441 Hz from 0-1 sec and f = 882 Hz from 1-2 sec), apply your track_pitch_fftmax(), (blockSize = 1024, hopSize = 512) and plot the f0 curve. Also, plot the absolute error per block and discuss the possible causes for the deviation. Repeat for track_pitch_hps() with the same signal and parameters. Why does the HPS method fail with this signal?

This signal is composed of two sinusoids with no harmonics. Thus, an analysis method based on harmonics will fail.

E2. Next use (blockSize = 2048, hopSize = 512) and repeat the above experiment (only for the max spectra method). Do you see any improvement in performance?

The error with block size 1024 is about 10 Hz for the first second and 20 seconds for the second second. The error with blocksize 2048 is the same for the first second, but zero for the second second.

E3. Evaluate your track_pitch_fftmax() using the development set Download development set(see assignment 1) and the eval_pitchtrack_v2() method (use blockSize = 1024, hopSize = 512). Report the average performance metrics across the development set.

Average RMS: 1922.223149488458
Average PFP: 81.73189199051268
Average PFN: 0.06018375606660836

E4. Evaluate your track_pitch_hps() using the development set and the eval_pitchtrack_v2() method (use blockSize = 1024, hopSize = 512). Report the average performance metrics across the development set.

Average RMS: 752.5711452792799
Average PFP: 13.097746761539865
Average PFN: 24.865401357193765

E6. Evaluate your track_pitch() using the development set and the eval_pitchtrack_v2() method (use blockSize = 1024, hopSize = 512) over all 3 pitch trackers (acf, max and hps) and report the results with two values of threshold (threshold = -40, -20)

Tracking pitch for developmentSet/trainData/01-D_AMairena.wav using acf and threshold of -40
RMS: 204.820710036163
PFP: 70.83333333333334
PFN: 0.05040322580645161

Tracking pitch for developmentSet/trainData/24-M1_AMairena-Martinete.wav using acf and threshold of -40
RMS: 623.2211589876858
PFP: 77.77777777777779
PFN: 0.0

Tracking pitch for developmentSet/trainData/63-M2_AMairena.wav using acf and threshold of -40
RMS: 352.187481364272
PFP: 91.37931034482759
PFN: 0.0

Tracking pitch for developmentSet/trainData/01-D_AMairena.wav using acf and threshold of -20
RMS: 208.84846242044927
PFP: 100.0
PFN: 0.0

Tracking pitch for developmentSet/trainData/24-M1_AMairena-Martinete.wav using acf and threshold of -20
RMS: 623.2211589876858
PFP: 100.0
PFN: 0.0

Tracking pitch for developmentSet/trainData/63-M2_AMairena.wav using acf and threshold of -20
RMS: 352.187481364272
PFP: 100.0
PFN: 0.0

Tracking pitch for developmentSet/trainData/01-D_AMairena.wav using max and threshold of -40

RMS: 1694.486662346608
PFP: 59.583333333333336
PFN: 0.15120967741935484

Tracking pitch for developmentSet/trainData/24-M1_AMairena-Martinete.wav using max and threshold of -40
RMS: 1988.9020509931695
PFP: 72.22222222222221
PFN: 0.0

Tracking pitch for developmentSet/trainData/63-M2_AMairena.wav using max and threshold of -40
RMS: 2083.527940550901
PFP: 62.643678160919535
PFN: 0.07974481658692185

Tracking pitch for developmentSet/trainData/01-D_AMairena.wav using max and threshold of -20
RMS: 1694.2394569213031
PFP: 84.16666666666667
PFN: 0.10080645161290322

Tracking pitch for developmentSet/trainData/24-M1_AMairena-Martinete.wav using max and threshold of -20
RMS: 1988.9020509931695
PFP: 92.06349206349206
PFN: 0.0

Tracking pitch for developmentSet/trainData/63-M2_AMairena.wav using max and threshold of -20
RMS: 2083.527940550901
PFP: 68.96551724137932
PFN: 0.07974481658692185

Tracking pitch for developmentSet/trainData/01-D_AMairena.wav using hps and threshold of -40
RMS: 597.4478543084341
PFP: 13.750000000000002
PFN: 24.04233870967742

Tracking pitch for developmentSet/trainData/24-M1_AMairena-Martinete.wav using hps and threshold of -40
RMS: 843.1149761486033
PFP: 10.317460317460316
PFN: 14.719101123595504

Tracking pitch for developmentSet/trainData/63-M2_AMairena.wav using hps and threshold of -40
RMS: 796.7849377009161
PFP: 9.770114942528735
PFN: 35.88516746411483

Tracking pitch for developmentSet/trainData/01-D_AMairena.wav using hps and threshold of -20
RMS: 617.8135219883206
PFP: 16.25
PFN: 23.991935483870968

Tracking pitch for developmentSet/trainData/24-M1_AMairena-Martinete.wav using hps and threshold of -20
RMS: 843.1149761486033
PFP: 12.698412698412698
PFN: 14.719101123595504

Tracking pitch for developmentSet/trainData/63-M2_AMairena.wav using hps and threshold of -20
RMS: 796.7849377009161
PFP: 10.344827586206897
PFN: 35.88516746411483