ELECTRICAL AND ELECTRONICS ENGINEERING INSTITUTE

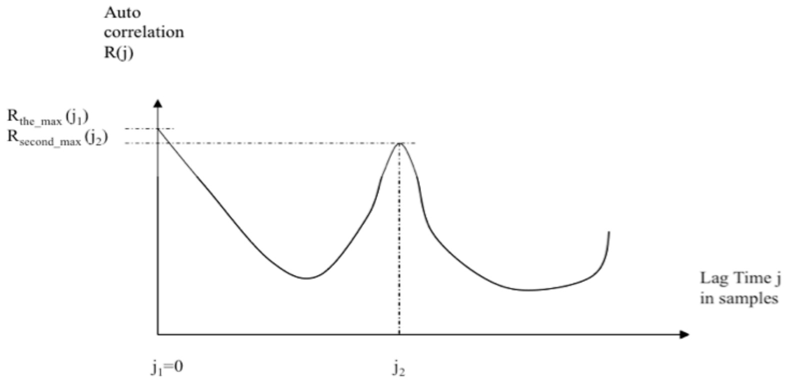
EE 286: Digital Audio Signal Processing

Exercise 5: Pitch Estimation

Prerequisites: Matlab, MIRtoolbox
https://www.mathworks.com/matlabcentral/fileexchange/24583-mirtoolbox

Pitch estimation is essential to many music signal processing applications such as genre classification, music style analysis, automatic transcription, etc. The objective in this exercise is to apply signal processing techniques to obtain an estimate of the pitch of an audio signal.

The main operation used in pitch estimation is autocorrelation. When a segment of a signal is correlated with itself, the distance between the positions of the maximum and the second maximum is defined as the fundamental period. The reciprocal of the fundamental period is the fundamental frequency or pitch.



1. Signal pre-processing
   Load the audio file *ragtime.wav* using the function wavread().
   Convert the stereo signal into mono signal by adding half the magnitude from each channel.
   Normalize the audio signal such that the maximum is 1 and the average is 0 using the equation
   $$x_{norm} = \frac{x - \bar{x}}{\max\left(|x - \bar{x}|\right)}$$

   Use the following specifications: FS=44.1 kHz, frame size = 30 msec, hop size = 20 msec. The frame size defines the size of the audio segment for analysis while the hop size defines the amount of time the analysis window moves. Given the specifications, how many frames will be analyzed in the audio file?

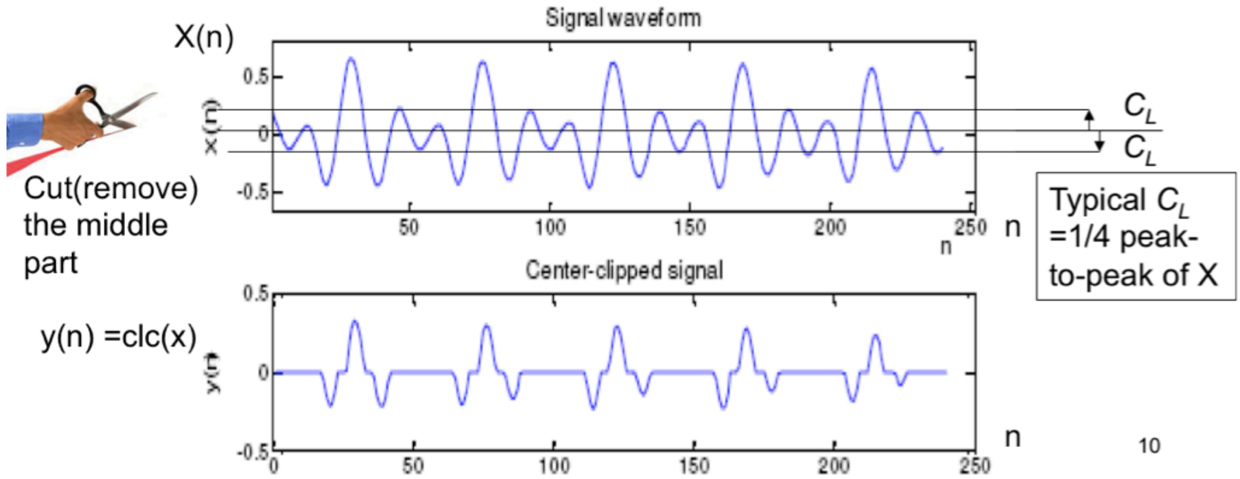
2. Autocorrelation
   Use the Matlab function xcorr() to compute the autocorrelation of a segmented audio. Note that the output of autocorrelation is even symmetric. Determine the fundamental frequency per analysis frame using the definition described previously. Hint: you may use miraudio(), mirpeaks() and mirgetdata() functions. Plot the results where the x-axis and y-axis correspond to time and fundamental frequency respectively.

3. Autocorrelation with center clipping

One variation of pitch estimation is autocorrelation with center clipping. The idea is that it may give more accurate results since higher frequencies will not interfere.

$$y(n) = clc[x(n)] = \begin{cases} (x(n) - C_L), & x(n) \geq C_L \\ 0 & , |x(n)| \langle C_L \\ (x(n) - C_L), & x(n) \leq C_L \end{cases}$$

$$R'(m) = \sum_{n=0}^{N-1-m} y(n) \cdot y(n+m), \; 0 \leq m \leq M_0$$



Set C_L to 0.1. Determine the fundamental frequencies and plot the results. How does this compare with the results in number 2? Change C_L to 0.2, how does this new value affect the results?

4. Effect of frame size

If the frame size and hop size are increased to 60 msec and 40 msec respectively, how do these changes affect the estimates in numbers 2 and 3? Plot the fundamental frequency estimates.

You can work as a pair. Together with your answers to the questions, submit the Matlab script yoursurname1_ yoursurname1_pitch.m. Compress your files as a zip file then upload in UVLE.

DEEE Building, Velasquez Street, University of the Philippines Diliman, Quezon City, 1101 Philippines
Tel: 981-8500 local 3300 TeleFAX: 925-2957 URL: http://www.up.edu.ph/~eee