Music 421A
Spring 2019-2020
Homework #7
FFT processing, Noise reduction
One week assignment

Note: If taking this course online and self-paced, the *Week 8* lecture videos contain more about overlap-save and (constant) overlap-add. Watching those now may be helpful for this assignment.

Theory Problems

1. (5 pts) **Optimal Window Size**
   We wish to implement FFT convolution to apply a fixed causal FIR filter $h(n)$ having length N_h. The input signal may be any length, say L, so we have to choose a finite frame size M and FFT size N. Assuming the cost of an FFT to be exactly $N \lg N$, and assuming a hop size $R = M$ (rectangular window as usual for FFT convolution), find the optimal window length M (a function of N_h) which minimizes computational cost.

2. (5 pts) **Overlap Save** versus **Overlap Add**
   Repeat the previous problem for the *Overlap-Save* method of FFT convolution. (You may look it up anywhere you wish, such as its Wikipedia page.)

3. (13 pts) **Spectral Smoothing**
   Suppose we have used the FFT to compute the spectrum $X(\omega_k)$, $k = 0, 1, \ldots, N-1$ of a length $M = N/2$ signal frame $x_m(n)$ centered at time mR samples using a length M (causal) rectangular window with unit amplitude, and suppose we now smooth this spectrum using the length N smoothing kernel $H = [1, 0, -1/2, 0, 0, \ldots, 0, 0, -1/2, 0]$.

   (a) (5 pts) What is the new effective window w_s in the time domain?

   (b) (5 pts) A window w and hop size R are said to be Constant-OverLap-Add (COLA) when

   $$\sum_{m=-\infty}^{\infty} w(n - mR)$$

   is constant for all n. In other words, the sum of overlapped windows is constant. Given that the maximum COLA hop size for the original rectangular window was M, find is the maximum COLA hop size for the new effective window w_s.

   (c) (3 pts) Verify your result by plotting the new effective window, overlapped and added to itself at the new maximum COLA hop size, using 5 copies of the window in the overlap-add (corresponding to 5 successive values of m).

Lab Assignment

1. (35 pts) **Bandpass Filtering for Noise Reduction and Dynamic Range Compression**
   Write a program to denoise the corrupted birdsong in `wrenpn1.wav`[1] by means of a time-varying FIR bandpass filter. Assume that the birdsong consists of a single frequency- and amplitude-modulated sinusoid. A dynamic range compression scheme will also be implemented. The guideline on how to do this is given below.

   (a) (5 pts) Read in frames of signal using a Hann window of length $M = 255$ and a hop-size which will give a constant overlap-add. *State clearly the hop-size you use.* Choose a reasonable FFT length N to use with this frame-size M and explain why you chose it.

   (b) (5 pts) For each frame of noisy signal, find the single largest peak of the spectrum using `findpeaks.m` that you made for the previous homework.

   (c) (5 pts) Design a narrow bandpass filter $H(\omega)$ centered at the frequency just obtained using the same window used in analysis with the same length, $L = M$. Be sure that the filter has a unity DC gain in order to preserve our sinusoid's amplitude. [Hint: Use the window method for FIR digital filter design.]

   (d) (5 pts) Multiply $H(\omega)$ and $X_w(\omega)$ to obtain the filtered output spectrum where $H(\omega)$ and $X_w(\omega)$ is the Fourier transform of the filter impulse response and the windowed signal respectively. Be sure that your zero-padding is enough to avoid time-aliasing. *State clearly what FFT length you use.*

   (e) (5 pts) IFFT the resulting spectrum and overlap-add the frame into the output buffer. Make sure that your overlap-add is correctly aligned.

   (f) (10 pts) *Dynamic range compression*: Equalize each output frame (before overlap-add) so that they are all of equal amplitude level. This can be done, for example, by normalizing each frame by its root-mean-square (RMS). Be sure that the final output is below saturation. However, disable the normalization for all frames having a small RMS value (e.g., near that of the noise during intervals of silence). One effective technique is to define a frame gain that is fixed for the frame, but varies smoothly from 1/rms to 0 as a function of signal-to-noise ratio.

   Submit the denoised output sound with and without dynamic range compression.

   Also submit

   (a) plots of the spectrograms of the denoised output with and without dynamic range compression.

   (b) a plot the magnitude spectrum of one frame of the output (before dynamic range compression) superimposed on that of the input at the time halfway through.

[1] `http://ccrma.stanford.edu/~jos/sasp/hw/wrenpn1.wav`