ECE 797: Speech and Audio Processing

Assignment #3: Homomorphic analysis and synthesis of speech

The Due date: Report and MATLAB code to be submitted by Thursday, April 8.

The report: Should include plots and descriptions of results and answers to specific questions. A hand-written copy of any mathematical derivations for Part A is acceptable.

The code: The MATLAB M-files generating the results in your report should be submitted.

The rules: Each student should work on their own code and report. If you get stuck, you can ask for advice from another student, but please do not show each other your mathematical derivations for Part A. If you get really stuck, you can ask me.

The Assignment:

Part A: Comparison of the spectral root and complex cepstrum.

1. For $x[n] = p[n] * h[n]$, with $p[n]$ a periodic impulse train, discuss the implications of $p[n]$ not being aligned with the origin for both the complex and spectral root cepstrum.
2. Using the relation between the spectral root cepstrum and the complex cepstrum given by Eq. (6.20) of Quatieri, show that as γ approaches zero, the spectral root deconvolution system approaches the complex cepstral deconvolution system.
3. Show why it is necessary to perform two phase-unwrapping operations, one for the forward $X(\omega)^{\gamma}$ and one for the inverse $X(\omega)^{1/\gamma}$ transformations when γ *is not an integer*. How does this change when γ *is an integer*?

Part B: Implementation of homomorphic analysis, filtering and synthesis in MATLAB.

In this part, use the file `assgn2.wav` that was used also in Assignment #2, which contains 250 ms of a voiced speech segment sampled at 10 kHz.

1. Apply a 25 ms Hamming window to the beginning of the speech segment. Using a 1024-point FFT, compute the real cepstrum of the windowed signal and plot it. For a clear view of the real cepstrum, set the first cepstral value to zero (which is the DC component of the log-spectral magnitude) and plot only the first 256 cepstral values.

2. Estimate the pitch period (in samples and in milliseconds) from the real cepstrum by locating a distinct peak in the quefrency region.

3. Extract the first 50 low-quefrency real cepstral values using a lifter of the form:

$$l[n] = \begin{cases} 1, & n = 0, \\ 2, & 1 \le n \le 49, \\ 0, & \text{otherwise,} \end{cases}$$

   then apply the inverse cepstral transform (using a 1024-point FFT again) to obtain a minimum-phase estimate of the vocal-tract impulse response. Plot just the first 200 samples to obtain a clear view. Does the impulse response resemble one period of the original waveform? If not, then why not?

4. Using your results from 2 & 3 above, synthesize a 250-ms estimate of the unwindowed speech segment from `assgn2.wav`, assuming that the excitation source is a perfectly period train of ideal impulses. How does your waveform estimate differ from the original? Consider the minimum-phase nature of the impulse response estimate.

5. Plot the log-magnitude and phase spectra of the impulse response estimate obtained in 3 above and compare them to the log-magnitude and phase spectra of the windowed speech segment.

6. Compare the results obtained in this assignment (for both the analysis and the synthesis parts) with the results you obtained using linear prediction analysis and synthesis in Assignment #2.