EE 114: Speech and Image Processing Systems Design Fall 2017

Midterm

November 6, 2017

- The total number of points available on this midterm is 100.
- In solving the problems on this midterm, clearly show your work and clearly indicate your answer. For all problems on this test, answers should be expressed in the simplest form possible.
- There are a total of 4 problems, each with several subparts. Please verify your exam to make sure you have the entire exam (total of 4 problems, total length 4 pages).
- $\log_{10}(2) \approx 0.3$
- An octave corresponds to a factor of 2 change in frequency. For example, given a tone at 200 Hertz, a tone an octave above this tone would have a frequency of 400 Hz. Similarly, a tone an octave below 200 Hz is at 100 Hz.

1. Discrete Time Fourier Transform (25 points)

Note: In all parts of this problem, sequences that are a function of n are defined only at integer values of n.

Let $x(n)$ be a signal with DTFT $X(\omega)$. In addition, let $v(n)$ be as follows:

$$v(n) = \frac{e^{-j\pi n}x(n) + x(n)}{2}$$

a) (8 points) Let $V(\omega)$ denote the DTFT of $v(n)$. Express $V(\omega)$ in terms of $X(\omega)$.

b) (9 points) Now consider the sequence $y(n) = v(2n)$. For example, if $v(n)$ is a sequence of the form

$\underline{a}$ b c d e f g...

then $y(n) = v(2n)$ would be

$\underline{a}$ c e g...

Find $Y(\omega)$, the DTFT of $y(n)$, expressed in relation to $V(\omega)$, using the most simplified form of this expression as possible. As noted in part a), $V(\omega)$ denotes the DTFT of $v(n)$.

c) (8 points) Let $g(n) = x(2n)$ and let G(ω) denote the DTFT of $g(n)$. Express G(ω) in terms of X(ω).

2. Sampling (22 points)

Consider a continunous time signal that is then sampled at a rate of 20.48 kHz. A segment of length 1024 samples is selected and the 1024-point DFT is computed.

a) (5 points) What is the time duration of the segment?

b) (5 points) What is the frequency resolution of the DFT, where for this problem frequency resolution is defined as the spacing, in Hertz, between two adjacent DFT bins?

c) (12 points) Now consider a continunous time signal signal that consists solely of a pure sinusoid at 15 KHz. Suppose that the sinusoidal signal is then sampled at 20.48 KHz. The sampling process does not include any filtering. Sampling in this case consists simply of measuring the exact value of the input signal at intervals of T seconds, where T is the inverse of the sampling frequency. Then, the sampled signal is reconverted to analog form by passing it through an ideal D/A converter. What frequency, in Hertz, will be output from the D/A converter?

3. Linear Predictive Coding and Autocorrelation (28 points)

Consider a vowel sound sampled at 8 KHz. A segment, denoted $s(n)$, is extracted using a Hamming window. The magnitude of the DTFT of this segment is shown in Figure 1 below (assume that in the figure all the harmonics shown are equally spaced):

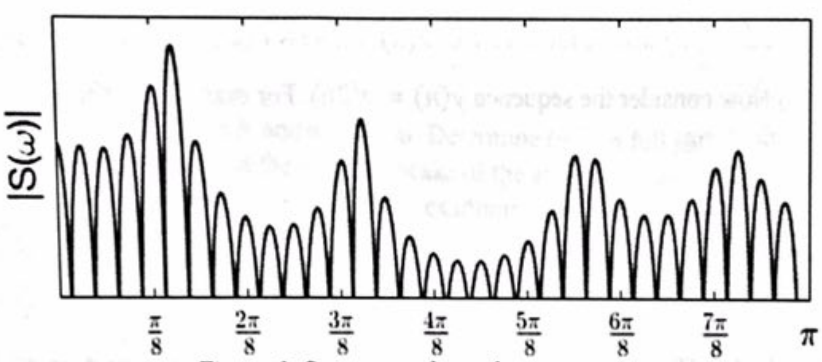

Figure 1. Spectrum of vowel segment s(n)

a) (8 points) Given the lobes shown in Figure 1 above, and the fact that this spectrum was obtained by taking the DTFT of a signal that had been subject to a Hamming window, what is the minimum window length that was needed to ensure that the lobes associated with the harmonics shown in the figure (which for each harmonic correspond to the result of convolving the DTFT of the window function with the delta function associated with that harmonic) remain non-overlapping? You should specify the window length in samples, not seconds.

For parts b) through d) of this problem, consider the following:

The time domain signal representing the vowel segment is passed through the filter denoted "System 1" as shown in Figure 2 below, and is then down-sampled to a sampling frequency of 1600 Hz. System 1 is a filter. The downsampler downsamples by a factor of 5. This means that it simply keeps every fifth sample and discards the intervening samples. For example, if the input to the downsampler is a sequence

1 2 3 4 5 6 7 8 9 10 11 12...

the output is a sequence

1 6 11...

While System 1 is a filter, there is no additional filter included in the box in Figure 2 labeled "Downsampler (Decimator)". The downsampled signal is then passed through an LPC predictor characterized by the following H(z).

$$H(z) = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$

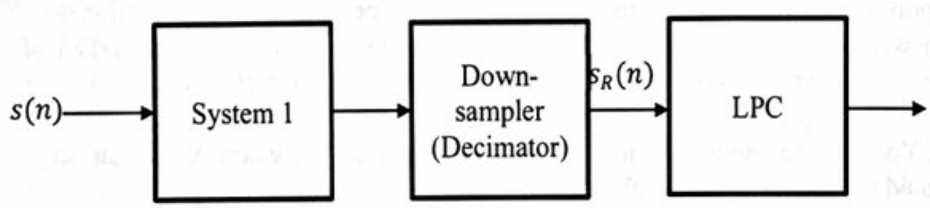

Figure 2.

b) (5 points) What filter should be used in System 1 to ensure that there is no aliasing arising due to the subsequent downsampling step in which only every fifth sample is kept? You should assume that the filter is ideal (i.e., the frequency response can contain vertical edges). Your answer should specify the type of filter as well as the specific attributes.

c) (10 points) Using the second-order linear predictive analysis, find the predication coefficients a_k, assuming that $R(0) = 6, R(1) = -4, R(2) = 1$.

d) (5 points) Using the result from part c), and assuming a sampling frequency of 1600 Hz, find the formant frequenc(y)(ies) in **Hertz** associated with the second order model associated with the R(0), R(1), and R(2) values provided in part c).

4. Windowing (25 points)

Consider a discrete-time triangular window $w_T(n)$ specified as follows, where N is even:

$$w_T(n) = \begin{cases} \dfrac{n}{N/2}, & n = 0, \dots, \dfrac{N}{2} \\ \dfrac{N-n}{N/2}, & n = \dfrac{N}{2} + 1, \dots, N-1 \end{cases}$$

Note that since N is even, the number of nonzero values in $w_T(n)$ will be $N - 1$, which is odd.

a) (15 points) Find $W_T(\omega)$, the DTFT of the window $w_T(n)$. Note that this problem is not asking you to find the absolute value of $W_T(\omega)$. It is asking you to find $W_T(\omega)$, which may be complex.

b) (10 points) Assume a large N and a small ω. Determine (with a full justification of your answer) the roll-off rate of the sidelobe peaks of the absolute value of the DTFT of the triangular window (again assuming you are examining the rolloff in where ω is small, and also assuming a large N). Your answer should be expressed in terms of in dB per octave. Your answer should be a number (e.g., X dB per octave, where X is a number), and should not contain any unevaluated log functions.