Department of Electrical and Computer Engineering
Digital Speech Processing
Homework No. 5

Problem 1

Figure 1 shows plots of 6 synthetic vowel short-time log magnitude spectra as obtained using a Hamming window of an appropriate length. The set of 6 spectra cover the vowel sounds /IY/ (as in /beet/), /AA/ as in /hot/, and /UW/ as in /boot/.

(a) What is the pitch of each of the 6 vowels?
(b) Which of the plots correspond to each of the 3 vowels?
(c) What happens to the spectral plots of some of the vowels at high frequencies? Why does this occur?

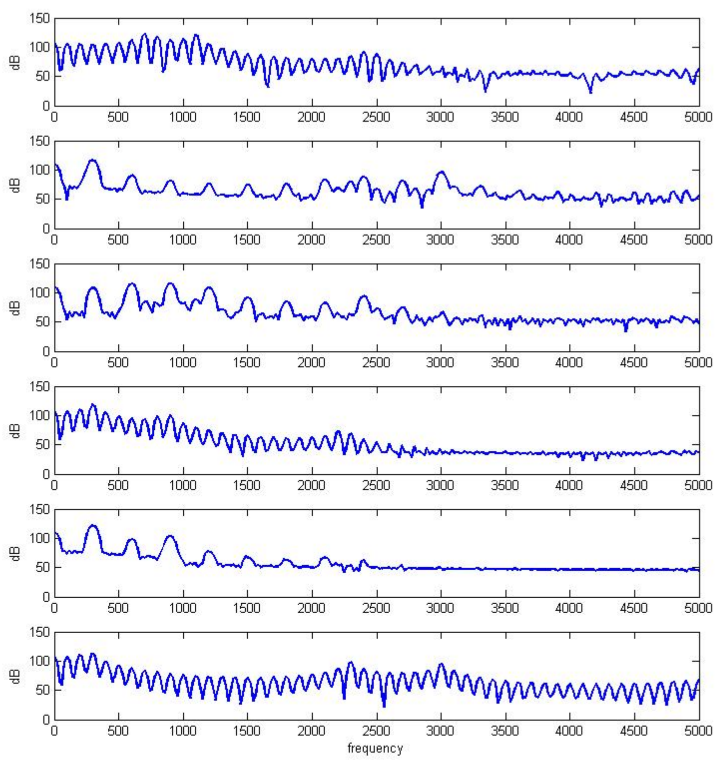

Figure 1: Synthetic Vowel Spectra

Problem 2

Figure 2 shows plots of 6 speech short-time log magnitude spectra as obtained using a Hamming window of an appropriate length. The set of 6 spectra include vowel and consonant regions by a male, a female and a child talker.

(a) Which of the 6 spectra are most likely to have been uttered by a child? What leads you to this conclusion?

(b) Which of the 6 spectra correspond to voiced sounds?

(c) Which of the voiced speech spectra most likely come from an adult male; which from an adult female?

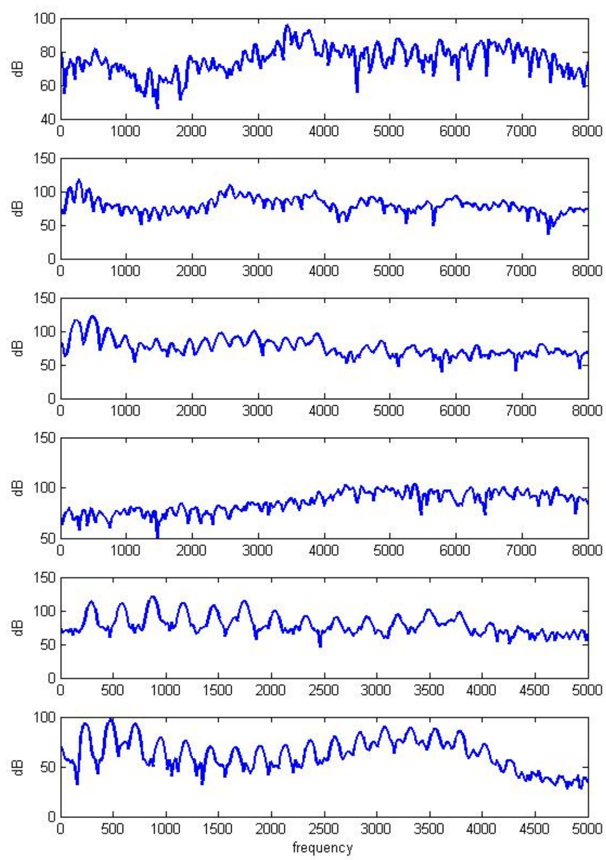

Figure 2: Speech Spectra

Problem 3

A digital signal, $x[n]$, is defined as:

$$x[n] = \begin{cases} 7 & n = 0 \\ 6 & n = 1 \\ 6 & n = 2 \\ 7 & n = 3 \\ 3 & n = 4 \\ 1 & n = 5 \\ 0 & n \geq 6, n \leq 0 \end{cases}$$

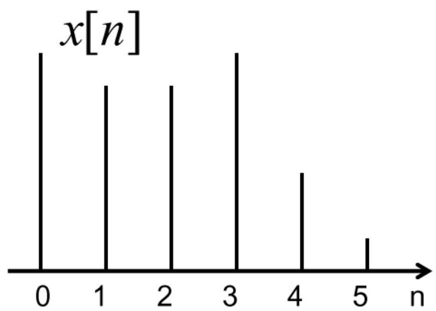

Figure 3: Plot of digital sequence.

The digital signal is plotted in Figure 3. $X(e^{j\omega})$, the discrete Fourier transform of $x[n]$, is sampled at N points around the unit circle, i.e., at the set of frequencies:

$$\omega_k = \frac{2\pi}{N}k, \qquad k = 0, 1, \ldots, N-1$$

and the resulting sequence, $\hat{X}[k]$, is obtained as:

$$\hat{X}[k] = X(e^{j\omega})|_{\omega=2\pi k/N}$$

Finally we inverse transform $\hat{X}[k]$ and obtain the sequence $\hat{x}[n]$. Determine and plot $\hat{x}[n]$ for $N = 40$, $N = 10$, $N = 5$, and $N = 3$.

Problem 4

A digital signal, $x[n]$, has the form:

$$x[n] = r^n u[n], \quad |r| < 1$$

(a) Solve for $X(e^{j\omega})$, the Fourier transform of $x[n]$.

(b) $X(e^{j\omega})$ is sampled at a set of N uniformly spaced points between $\omega = 0$ and $\omega = 2\pi$, giving the sampled signal, $\tilde{X}[k]$, which is formally defined as:

$$\tilde{X}[k] = X(e^{j\omega})|_{\omega=2\pi k/N}, \quad k = 0, 1, \ldots, N-1$$

Solve explicitly for $\tilde{x}[n]$, the inverse Fourier transform of $\tilde{X}[k]$.
(c) Plot $\tilde{x}[n]$ for $n = 0, 1, \ldots, 2N-1$ for the case where $N = 5$ and $r = 0.9$.

Problem 5

In implementing time-dependent Fourier representations, we employ sampling in both the time and frequency dimensions. In this problem we investigate the effects of both types of sampling.

Consider a sequence $x[n]$ with conventional Fourier transform

$$X(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x[m]e^{-j\omega m}$$

(a) If the periodic function $X(e^{j\omega})$ is sampled at frequencies $\omega_k = 2\pi k/N,\ k = 0, 1, \ldots, N-1$, we obtain

$$\tilde{X}[k] = \sum_{m=-\infty}^{\infty} x[m]e^{-j\frac{2\pi}{N}km}$$

These samples can be thought of as the discrete Fourier transform of the sequence $\tilde{x}[n]$ given by

$$\tilde{x}[n] = \frac{1}{N}\sum_{k=0}^{N-1} \tilde{X}[k]e^{j\frac{2\pi}{N}kn}$$

Show that

$$\tilde{x}[n] = \sum_{r=-\infty}^{\infty} x[n+rN]$$

(b) What are the conditions on $x[n]$ so that no aliasing distortion occurs in the time domain when $X(e^{j\omega})$ is sampled?

(c) Now consider "sampling" the sequence $x[n]$; i.e., let us form the new sequence

$$y[n] = x[nM]$$

consisting of every M^{th} sample of $x[n]$. Show that the Fourier transform of $y[n]$ is

$$Y(e^{j\omega}) = \frac{1}{M}\sum_{k=0}^{M-1} X(e^{j(\omega-2\pi k)/M})$$

In proving this result you may wish to begin by considering the sequence

$$v[n] = x[n]p[n]$$

where

$$p[n] = \sum_{r=-\infty}^{\infty} \delta[n + rM]$$

Then note that $y[n] = v[nM] = x[nM]$.

(d) What are the conditions on $X(e^{j\omega})$ so that no aliasing distortion in the frequency domain occurs when $x[n]$ is sampled?

Problem 6

Write a MATLAB program to display both a narrowband and a wideband spectrogram (on a gray scale) of a speech file. Explain your choices of spectrogram parameters (window, window duration, frame overlap). Test your MATLAB program on the file test_16k.wav and turn in both the program and the pair of spectrograms for this speech file, plotted on a single sheet of paper.

Problem 7

Write a MATLAB program to speed up a speech file by a factor of 2-to-1. Use the method of overlap-add to analyze the STFT of the signal (using a rectangular window of length 512 samples with 256 sample overlap between frames), throw out every other frame, and resynthesize the speeded-up speech. Plot the original speech file and the speeded up speech file, and plot the narrow band spectrograms of both the original and the speeded up speech files. Use the speech file test_16k.wav to test your program.