EECS 225d: Audio Signal Processing by Humans and Machines

Revised Problem Set 1
Due in class Wed Feb 5, 1997

READING: Chapters 1-5

1. Find the Voder sequence for any of the practice sentences of Fig 2.8 (Lesson 37). Break the sentence into phonemes shown on CLASSIFICATION OF THE SPEECH SOUNDS handout and list with voicing (wrist bar) as well as console keys used. Note that BK1, BK2, and BK3 keys in the handout are the k-g, p-b, and t-d keys of the text's figure 2.5. For the "Combinational and Transitional Sounds" in section 2 of the handout, you needn't specify the voder keys (unless you happen to know them off hand...)

   VODER EXAMPLE: "The voder can speak well."

/th/	Voiced	10Q
/ŭ/	Voiced	3458
/v/	Voiced	67Q
/o/	Voiced	3–2
/d/	Voiced	BK3
/r/	Voiced	36
/k/	Unvoiced	BK1
/á/	Voiced	457
/n/	Voiced	1
/s/	Unvoiced	9
/p/	Unvoiced	BK2
/e/	Voiced	18
/k/	Unvoiced	BK1
/w/	- - - - -	- -
/e/	Unvoiced	37
/ł/	Voiced	2

2. In lecture 2, we stressed the difference between articulatory-based synthesis and auditory-based synthesis. Categorize each of the following as either articulatory-based or auditory-based. Briefly discuss.
   (a) Telharmonium (p. 15 of Chapter 2)
   (b) Wheatstone / Von Kempelen Speaking Machine
   (c) Theremin
   (d) Player Piano
3. Chapter 3, Exercise 1

   Note that wide-band and narrow-band spectrograms differ in the length of their analysis windows. Wide-band spectrograms typically analyze the signal in 3 ms chunks while narrow-band spectrograms will typically use 20 ms sections. These pieces are then fourier transformed and their log magnitudes are displayed in adjacent vertical slices. The sections are overlapped in time so that a continuous graph is seen. (I.e. the first piece might be from time 0..3ms, then 1..4ms, then 2..5ms, etc.)

   HINT: Typical human speech has a pitch ranging from about 50 Hz for low voices to 600 Hz for high female voices and 800 Hz for children. Assume a pitch of 50-400 Hz is the typical range while people with voices with pitches in the low part of this range make understanding this problem easier!
4. Chapter 3, Exercise 4
5. Chapter 3, Exercise 5. Otherwise stated as:

   Construct a table for the phonemes of the phrase "We pledge you some heavy treasure" as shown in figures 3.6 - 3.8. For each of the phonemes, list your best estimate estimate of the start and end time. Do the same for the transition periods between the phonemes. (Note: cs stands for centiseconds as used in figures 3.6 - 3.8)

	/w/	1 cs	-	9 cs
E.g.:	/w/ → /E/	9 cs	-	11 cs
	/E/	11 cs	-	27 cs
6. Chapter 4, Exercise 1
7. Chapter 4, Exercise 4
8. Chapter 5, Exercise 1
9. Chapter 5, Exercise 2
10. Chapter 5, Exercise 3