Signal Processing
Lab 3: Speech synthesis

Ahmed Besbes *Panagiotis Koutsourakis* *Loic Simon*
Chaohui Wang

November 20, 2008

1 A simple model for producing vowels

In this lab you will implement a simple system that produces vowels, i.e. sounds like /A/, /o/, and /E/. This model is shown in figure **??**.

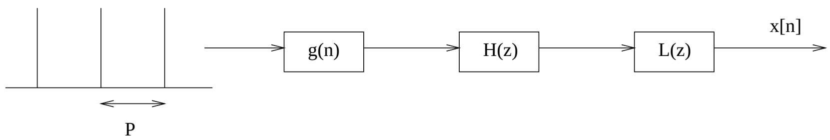

Figure 1: A simple model for the production of vowels

Vowels can be modeled as periodic signals. The male voice, on average, has a frequency of 100 Hz, while the female voice has a frequency of 200 Hz.

As shown on figure **??**, synthetic vowels are produced by a impulse train of period P samples

$$e[n] = \sum_{k=-\infty}^{\infty} \delta[n - kP] \tag{1}$$

filtered by a series of three filters $g(n)$, $H(z)$ and $L(z)$.

Q. Write a function to generate $e[n]$

- use prototype: `e = genpulse(f0,d,fs)` where `f0` is the frequency of the impulses, `d` is the duration of the signal and `fs` is the sampling rate (use 8000 Hz).

The first filter $g[n]$ has the impulse response:

$$g[n] = \begin{cases} \frac{1}{2}\left(1 - \cos\left(\pi n/N_1\right)\right) & 0 \leq n \leq N_1 \\ \cos(\pi(n - N_1)/(2N_2) & N_1) \leq n \leq N_1 + N_2 \\ 0 & \text{elsewhere} \end{cases} \tag{2}$$

(Rosenberg model of the glottal waveform for speech signals generation)

This filter corresponds to the vocal chords and its output is passed through the oral cavity.

Q. Write a function to generate $g[n]$

- use prototype: `g = rosenmodel(t1,t2,fs)` where `t1` and `t2` are durations corresponding to N_1 and N_2
- plot the impulse response and the output of the filter for `t1` $= 7.5$ms and `t2` $= 1.3$ms.
- use Matlab's `conv` to filter $e[n]$ with $g[n]$ (type `help conv` for more information)
- what is the main effect of $g[n]$?

The oral cavity is modeled by the filter $H(z)$. The oral cavity modulates the signal according to the sound we want to express and finally its output is modulated by our lips, that we model by the filter:

$$L(z) = 1 - z^{-1} \tag{3}$$

Q. What is the effect of filter $L(z)$?

- Give the expression of the transfert function of $L(z)$. What are its main properties ?
- tip: use Matlab function `freqz` to plot the transfer function (type `help freqz` for more information).

The system $H(z)$ changes according to the sound we want to produce. For instance, if the we want to pronounce the sound /A/, the spectrum of the filter is (on average) the one shown in figure **??**. Similarly the spectra of the filters for /o/ and /E/ are shown in figures **??** and **??** respectively.

Q. Write a function to synthetize filter $H(z)$

- read the help for Matlab function `filter`
- $H(z)$ should be designed using only poles. Thus, our function will generate the array corresponding to `A` in `help filter`.
- Since we want a real filter, all the poles will go by conjugate pairs:

  $$p = a_p e^{j2\pi f_p/f_s}, \;\; \bar{p} = a_p e^{-j2\pi f_p/f_s}$$

  for pole p with amplitude a_p and frequency f_p. Filter $H(z)$ will look like this:

  $$H(z) = \frac{1}{\prod_p \left(1 - pz^{-1}\right)\left(1 - \bar{p}z^{-1}\right)}$$

- use prototype `A = makevowel(a,f,fs)`, where `a` is an array of pole amplitudes and `f` is an array of frequencies (of same length).
- tip: one way to compute

  $$(a_1 - a_2 x)(b_1 - b_2 x) = a_1 b_1 - (a_1 b_2 + a_2 b_1)\, x + a_2 b_2 x^2$$

  with Matlab is to use `conv`. Indeed:

  $$[a_1 b_1, -(a_1 b_2 + a_2 b_1), a_2 b_2] = \texttt{conv}\left([a_1, -a_2], [b_1, -b_2]\right)$$

Construct the sounds /A/, /o/, /E/.

- use the information provided by the figures to chose the right frequencies for $H(z)$. The amplitude of the peaks will be determined manualy.
- you can use `soundsc` or `wavread` to listen to the sounds you created. Normalize the final signals so that the maximum absolute value is 1.
- do not forget to filter also by $L(z)$.
- concatenate the above sounds for various periods and listen to them all together.
- visualize the concatenated sound with the spectrogram tool from the previous lab (for instance, with command `figure, lab_spectrogram(sound, 100, fs, 4000)`).

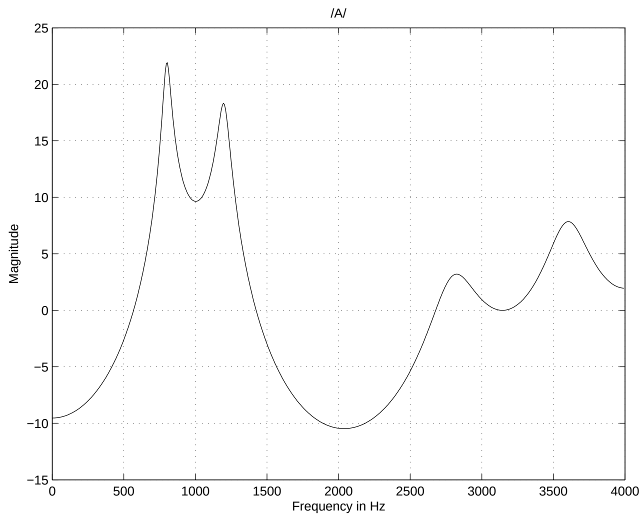

Figure 2: The spectrum of the filter of the oral cavity for the sound /A/ (Magnitude in log)

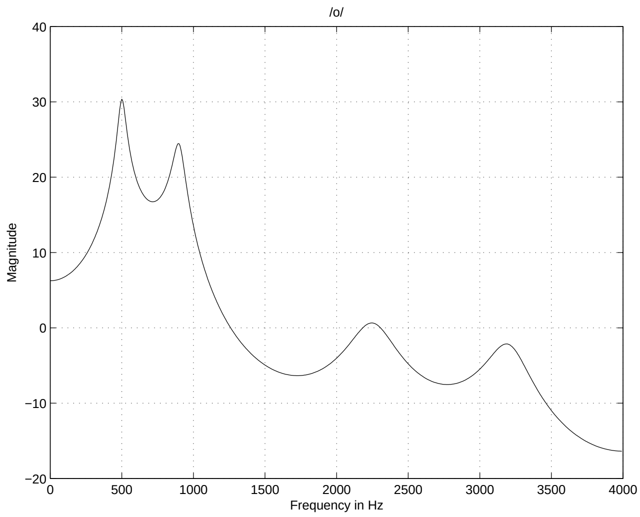

Figure 3: The spectrum of the filter of the oral cavity for the sound /o/

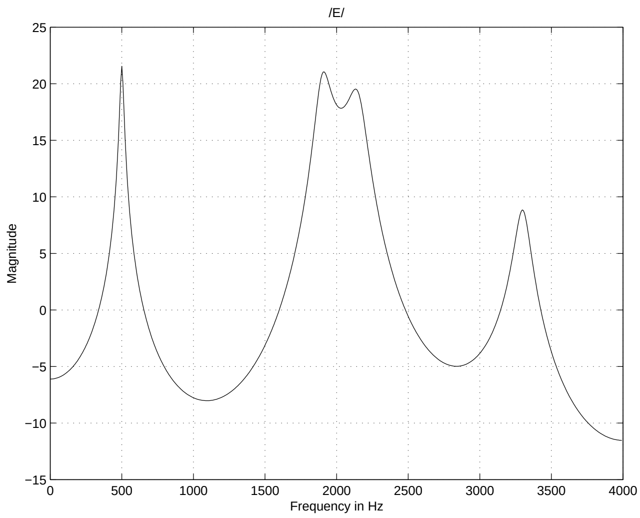

Figure 4: The spectrum of the filter of the oral cavity for the sound /E/