UNIVERSITY OF MIAMI
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
EEN 540 – DIGITAL SPEECH AND AUDIO PROCESSING

Project No. 2

Sound Production Modeling Using Concatenated Acoustic Tubes

Due date: Monday, March 11, 2009

This project will experiment with the production of sound through the vocal tract based on measurements of the vocal tract cross-sectional area, and sound filtering through an arbitrary tube.

The vocal tract will be approximated as a set of concatenated lossless tubes. Data on the area functions vowels /a/, /e/, /i/, /o/ and /u/ of a male speaker will be provided electronically. Female and child vocal tract lengths are approximately ¾ and ½ the male length respectively, with corresponding shapes extrapolated from male data. Losses in the system will be controlled with the introduction of a fictitious extra tube at the labial end. The sampling frequency for this study is 11025 Hz. Male students will deal with male and child speech; female students will deal with female and child speech.

1. Use the provided cross-sectional area values to determine the required minimum number of tubes in each vowel model. Plot the resulting area function, A(x) and reflection coefficients for each vowel. Note that the vocal tract length changes from vowel to vowel. Include the fictitious extra tube for the losses. See Fig 4.19 (a) and (b) of the textbook for an example of what you require to show.
2. The general form of the volume velocity transfer function, $V(z) = U_L(z)/U_G(z)$, is given in page 147 of the textbook. Utilize the iterative procedure outlined there to derive the transfer function of the system. Plot the magnitude response of $V(e^{j\omega})$ in dB scale for each of the five given vowels, both for lossless and (reasonably) lossy conditions. See Fig 4.19 (c) of the textbook for example.
3. Include the radiation model, $R_L(z) = (1-\alpha z^{-1})$, to compute and plot the magnitude response relating pressure at the lips to volume velocity at the glottis, i.e., $H(z) = AV(z)R_L(z)$.
4. Construct a reasonable glottal excitation signal, g(n), using the Rosenberg pulse. Plot 6 periods at a reasonable pitch frequency, F0, and the resulting magnitude spectrum of this signal.
5. The discrete-time terminal analog model, shown in Fig.4.20, will be used to produce speech. Filter the derived glottal excitation signal, g(n), with H(z) to generate speech, s(n), for each vowel. Produce 6 periods of output speech (pressure) signal and plot its time waveform and magnitude spectrum.
6. For each vowel vocal tract geometry generate 2 sec of speech and listen to the produced signal. Can you recognize the vowels? Do they sound natural? Experiment with parameter selection in order to improve the quality. Report on the best configuration.
7. Repeat parts 1 to 6 for child speech.

8. Suppose you are standing at the entrance of a 10-meter long tunnel of arbitrary shape (but no side cavities) reading out a couple of sentences. Plot the cross-sectional area of your tunnel and its concatenated tube approximation. Compute the transfer function of the tunnel and plot its magnitude response vs. frequency. Filter your input speech through the tunnel. Provide the input and output speech sounds. Repeat for a musical sound input.
9. Provide a critical discussion of your analysis and results.

2/25/2009 - MSS