

2.2 MATLAB Exercise – Three-Tube Vocal Tract Model

Program Directory: matlab_gui\Three_Tube_VT

Program Name: Three_Tube_VT_GUI25.m

GUI data file: Three_Tube_VT.mat

Callbacks file: Callbacks_Three_Tube_GUI25.m

TADSP: Section 5.2, pp. 200-208

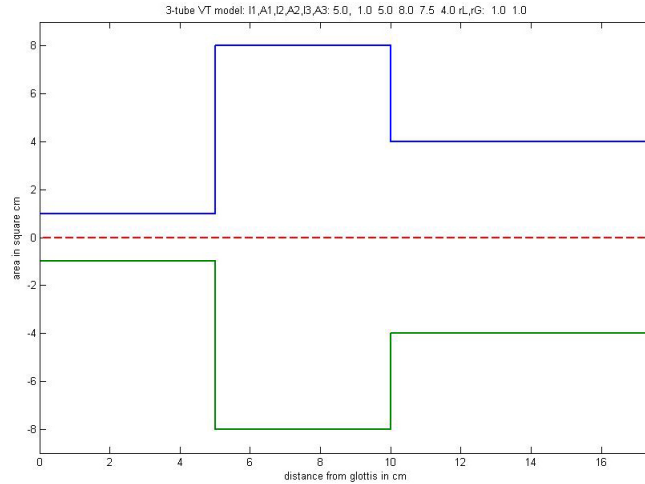


Figure 1: Three-tube model of human vocal tract.

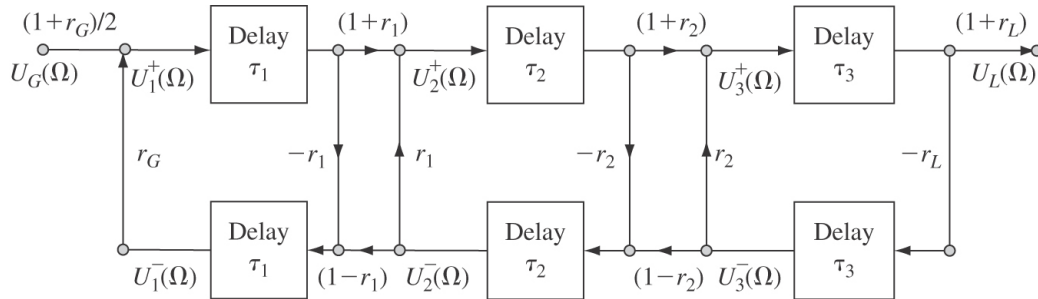


Figure 2: Complete flow diagram of a three-tube model of the vocal tract.

In this MATLAB Exercise¹ the user specifies the parameters of a lossless, three-tube model of the vocal tract (i.e., the parameters of the model include the length and area of the three tubes, along with glottal and lip reflection coefficients), and the exercise computes and plots the transfer function of volume velocity from the glottis to the lips. Such a three-tube model is shown in Figure 1, and the flow diagram of the composite system from the glottis to the lips is shown in Figure 2. The exercise computes and plots the volume velocity transfer function (log magnitude spectrum) from the glottis to the lips and determines the locations of the formant frequencies and bandwidths and prints them to a designated ascii text file. The 3-tube model frequency spectrum is sampled and converted to a 3-tube model impulse response which is convolved with a periodic impulse train and played out to the user so that the user can hear which vowel sound is produced.

¹This MATLAB exercise is virtually identical to the two-tube vocal tract model exercise with the difference being that the vocal tract is represented by a three-tube model, rather than a two-tube model.

Three-Tube Vocal Tract Model – Theory of Operation

By carefully solving the nodal equations of Figure 2, in terms of the forward and backward traveling waves from the glottis to the lips, assuming an input volume velocity of $U_G(\Omega)$, at the glottis, and an output of $U_L(\Omega)$, at the lips, the exercise computes the three-tube transfer function, giving the following:

$$V_a(\Omega) = \frac{U_L(\Omega)}{U_G(\Omega)} \quad (1)$$

$$= \frac{0.5(1 + r_L)(1 + r_1)(1 + r_2)(1 + r_G)e^{-j\Omega(\tau_1 + \tau_2 + \tau_3)}}{D(\Omega)} \quad (2)$$

where

$$D(\Omega) = 1 + r_2 r_L e^{-2j\Omega\tau_3} + r_1 r_L e^{-2j\Omega(\tau_2 + \tau_3)} + r_1 r_2 e^{-2j\Omega\tau_2} \\ + r_L r_G e^{-2j\Omega(\tau_1 + \tau_2 + \tau_3)} + r_2 r_G e^{-2j\Omega(\tau_1 + \tau_2)} \\ + r_1 r_G e^{-2j\Omega\tau_1} + r_1 r_2 r_L r_G e^{-2j\Omega(\tau_1 + \tau_3)}. \quad (3)$$

and where the lengths of the three tubes are l_1 , l_2 and l_3 , the areas of the three tubes are A_1 , A_2 , and A_3 , the reflection coefficients between tube sections are:

$$r_1 = \frac{A_2 - A_1}{A_2 + A_1} \quad (4)$$

$$r_2 = \frac{A_3 - A_2}{A_3 + A_2} \quad (5)$$

with the reflection coefficient at the glottis of r_G and the reflection coefficient at the lips of r_L , and the tube section delays are:

$$\tau_1 = l_1/c \quad (6)$$

$$\tau_2 = l_2/c \quad (7)$$

$$\tau_3 = l_3/c \quad (8)$$

The MATLAB Exercise uses the relation $\Omega = 2\pi f$ in order to evaluate Eqs. (2) and (3) for a dense grid of values of f , (in this case we use $f = 0, 1, \dots, 5000$ Hz). Next the exercise plots the log magnitude response of the three-tube model transfer function, and solves for the formants as the locations of the maxima of the log magnitude response. Furthermore the formant bandwidths are computed by finding the pair of frequencies (one above the formant and one below the formant) where the log magnitude response is a fixed threshold (`peak1`) below the log magnitude peak at the formant frequency. (In this case we used a value of 6 dB for `peak1`.) The locations of the formants are marked on the log magnitude plots. Furthermore, the log magnitude spectral calculation is carried out for a range of values for r_G and r_L since the bandwidth calculation is slightly dependent on values chosen for these two reflection coefficients. (The reader will recognize that it is impossible to determine the formant bandwidths when both r_G and r_L are set to 1 (i.e., a lossless vocal tract).)

Three-Tube Vocal Tract – GUI Design

The GUI for this exercise consists of two panels, 2 graphics panels, 1 title box and 16 buttons. The functionality of the two panels is:

1. one panel for the graphics display,
2. one panel for parameters related to specifying lengths and areas of the three-tube vocal tract model, and for running the program.

The pair of graphics panels is used to display the vocal tract shape (lower panel) and the log magnitude of the volume velocity transfer function from the glottis to the lips (upper panel). For all 'New Runs' the locations of the formants are also displayed on the upper graphics panel. The title box displays information about the lengths and areas of the three tubes that compromise the vocal tract model. The functionality of the 16 buttons is:

1. a text button that displays the current value of r_G ,
2. a text button that displays the current value of r_L ,
3. a slider button that allows the user to change the value of r_G in increments of 0.02 from its current value, when clicked on the slider button endpoints,
4. a slider button that allows the user to change the value of r_L in increments of 0.02 from its current value, when clicked on the slider button endpoints,
5. an editable button that specifies the length of the first vocal tract tube (in cm),
6. an editable button that specifies the length of the second vocal tract tube (in cm),
7. an editable button that specifies the length of the third vocal tract tube (in cm),
8. an editable button that specifies the area of the first vocal tract tube (in cm^2),
9. an editable button that specifies the area of the second vocal tract tube (in cm^2),
10. an editable button that specifies the area of the third vocal tract tube (in cm^2),
11. a pushbutton that starts a new run (a new set of values of any or all parameters, including r_G , r_L , 3-tube lengths and areas) to calculate and plot the volume velocity transfer function from the glottis to the lips, using the current values of the three-tube parameters,
12. a pushbutton that appends a new run (a new set of values of one or more of the tube parameters to the existing graphics plot (using a different color to display the new log magnitude spectrum). In this manner, multiple runs can be displayed on a single graphics plot,
13. an editable button that specifies the name of an ascii text file in which the formant center frequencies and bandwidths are written for the current (and all appended) vocal tract model run,
14. an editable button to specify the excitation period (in samples at a sampling rate of $f_s = 10000$ samples/second) of a pulse train that excites the 3-tube vocal tract impulse response and is played out to the user,
15. a pushbutton to play out the periodic vocal tract response created by exciting the 3-tube model impulse response by the periodic impulse train, of period ipd at $f_s = 10000$ Hz,
16. a pushbutton to close the GUI.

Three-Tube Vocal Tract Model – Scripted Run

A scripted run of the program 'Three_Tube_VT_GUI25.m' is as follows:

1. run the program 'Three_Tube_VT_GUI25.m' from the directory 'matlab_gui\Three_Tube_VT',
2. using the edit and slider buttons, adjust the three-tube model parameters. These parameters include r_G (the glottal reflection coefficient which is controlled by the first slider button and constrained to the range $-1 \leq r_G \leq +1$), r_L (the lips reflection coefficient which is controlled by the second slider button and again constrained to the range $-1 \leq r_L \leq +1$), l_1 , the length of the first tube in cm, l_2 , the length of the second tube in cm, l_3 , the length of the third tube in cm, A_1 , the area of the first tube in cm^2 , A_2 , the area of the second tube in cm^2 , A_3 , the area of the third tube in cm^2 , and the output text file whose default is the file `out_3_tube.txt`,

3. hit the 'New Run Three-Tube Model' button to begin a new run. The program plots the area of the three-tube model versus distance from the glottis in the lower graphics panel; the program also calculates and plots (in the upper graphics panel) the three-tube log magnitude frequency response. Finally, the program determines the resonance frequencies of the tube and estimates, wherever possible, the bandwidth of the resonances and prints the results in the Output Text File and overlays the resonance center frequency values on the log magnitude spectrum in the upper graphics panel,
4. the log magnitude frequency response for a new run of the three-tube vocal tract model can be obtained by first moving the sliders to change the values of r_G and r_L , and by manually changing the tube lengths and areas, and then pushing the 'Append Run' pushbutton. There is no restriction as to how many runs can be appended,
5. the excitation period (in samples at $f_s=10000$ samples/second) is set using the editable button that specifies ipd,
6. hit the button 'Play Vocal Tract Signal' to play out the results of convolving a periodic excitation signal with the 3-tube vocal tract impulse response,
7. experiment with a range of values for the three-tube vocal tract model parameters, thereby modeling a range of sounds that the three-tube models are capable of producing,
8. hit the 'Close GUI' button to terminate the run.

It is a fairly simple exercise to turn the three-tube model into a one-tube (uniform VT) model by making the areas of tubes 1 and 2 and 3 be the same. If you then change the lengths of the three tubes to add up to 17.5 cm, the resonances will be those of a uniform tube, namely 500, 1500, 2500, 3500, ... Hz. The exercise can also easily be changed to a two-tube model equivalent by making the areas of tubes 1 and 2 (or equivalently tubes 2 and 3) be the same, thereby creating a two-tube model. If you then change the lengths and areas of the three tubes to match configurations used for the two-tube exercise above, you can provide a sanity check on the code since it must match the results from the two-tube model exercise for these conditions.

The resulting graphical display for a three-tube model with lengths and areas of $l_1 = 8$ cm, $l_2 = 4$ cm, $l_3 = 5.5$, $A_1 = 8$ cm², $A_2 = 1$ cm², $A_3 = 8$ cm², and for a range of values of r_G and r_L is shown in Figure 3. This plot uses the 'Append New' pushbutton for all runs, and plots the results for values of r_G and r_L (with $r_G = r_L$) from 1.0 down to 0.7 in increments of -0.1 .

Three-Tube Vocal Tract Model – Issues for Experimentation

1. run the scripted exercise above, and answer the following:
 - what are the formant frequencies and bandwidths for the specified three-tube vocal tract model with $r_G = r_L = 1.0$?
 - what happens to the formant frequencies and bandwidths as r_L becomes less than +1, e.g., $r_L = 0.9$, and $r_G = 1.0$?
2. change the area of the third tube to 1.0; i.e., the same area as the second tube.
 - what type of vocal tract is described by this three-tube vocal tract model?
 - what are the locations of the formant frequencies for this three-tube vocal tract model?
 - how do the locations of the formant frequencies for this three-tube vocal tract model compare to those of the default two-tube vocal tract model of the previous MATLAB exercise?
3. using the default three-tube vocal tract model, change the area of the second tube to $A_2 = 8$
 - what are the new formant frequencies for this three-tube vocal tract model?
 - what type of vocal-tract model does this correspond to?

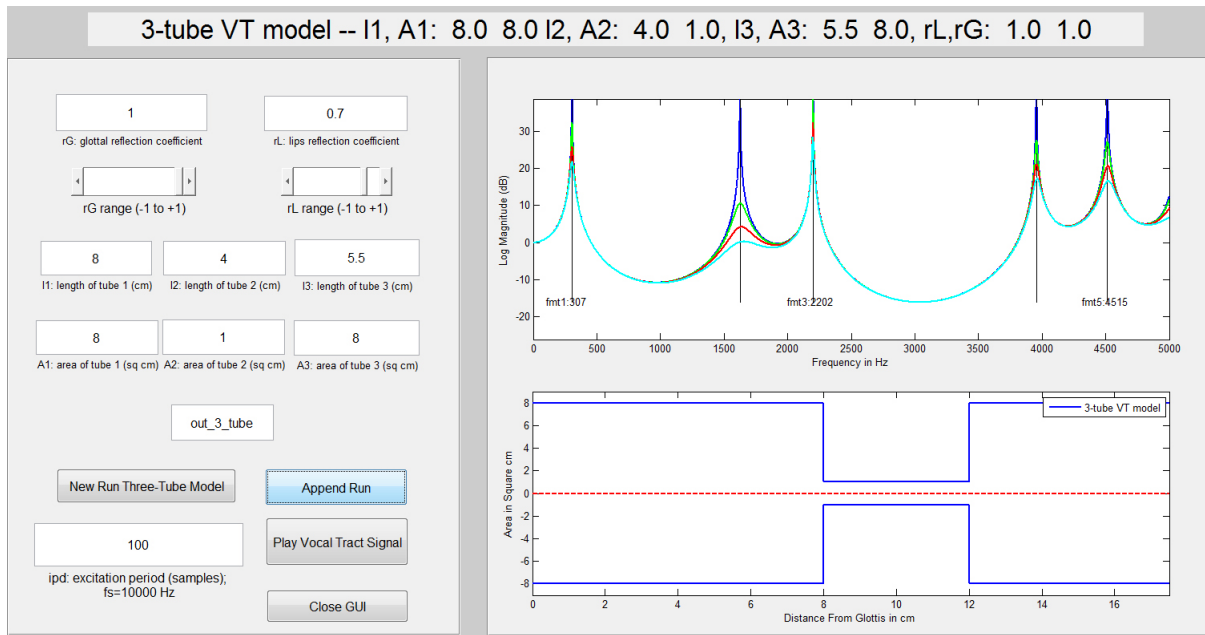


Figure 3: Three-tube vocal tract and the resulting volume velocity transfer function from the glottis to the lips for a fixed vocal tract shape, and a range of values of r_G and r_L .

1. choose areas and lengths to approximate a specified vowel by looking at figures in the textbook for the different vowel profiles.
2. do a run with parameters set for a total vocal tract length of 17.5 cm. Then shorten the tube lengths proportionally to see the effect of smaller vocal tract length.
3. run the program with a specified set of areas. Now scale each area by a constant and append the plot to show that it is the ratio of areas that controls the frequency response.
4. run the program with a specified set of areas and terminations. Now change one of the terminations and hit the 'Append Run' button to observe the effect on the formants.
5. run the program with a specified set of areas and terminations. Now change one of the tube lengths slightly and hit the 'Append Run' button to observe the effect on the formants.
6. run the program with a specified set of areas and terminations. Now change one of the tube areas slightly and hit the 'Append Run' button to observe the effect on the formants.