MATLAB Exercise – Vowel Synthesis

Program Directory: `matlab_gui\vowel_synthesis_rev1_gui25`
Program Name: `Vowel_Synthesis_GUI25.m`
GUI data file: `Vowel_Synthesis.mat`
Callbacks file: `Callbacks_Vowel_Synthesis_GUI25.m`
TADSP: Problem 13.3

This MATLAB exercise implements a simple vowel synthesis program.

Vowel Synthesis – Theory of Operation

This exercise synthesizes a vowel sound with a pitch frequency that linearly varies from a user-specified initial pitch frequency to a user-specified final pitch frequency. The formant bandwidths and center frequencies for each of 10 vowels are specified in the data file `vowels_fmts_bw.mat` along with the set of two-character ARPABet vowel names. Only three formant frequencies are specified for each vowel; it is assumed that the fourth resonance frequency is fixed at 4000 Hz for each vowel sound. The formant bandwidths for the first four formants are the same for each of the ten vowels and are set at the values 50, 80, 100, 150 Hz for the four bandwidths.

The synthesized vowel is played out to verify the correct operation of the code.

Vowel Synthesis – GUI Design

The GUI for this exercise consists of two panels, 8 graphics panels, 1 title box and 8 buttons. The functionality of the two panels is:

1. one panel for the graphics display,
2. one panel for parameters related to the synthesis of the variable pitch vowel sound, and for running the program.

The set of 8 graphics panels is used to display the following:

1. the pitch period contour for the vowel utterance,
2. the glottal pulse shape used for vowel synthesis,
3. the vowel sound obtained from convolving the pitch excitation signal with the glottal pulse and with the vowel impulse response based on four resonances,
4. the pitch excitation signal,
5. the vowel impulse response based on four resonances,
6. the vowel log magnitude response based on four resonances,
7. the log magnitude response of the final (convolved) vowel sound,
8. the linear magnitude response of the final (convolved vowel sound).

The title box displays the information about the selected vowel sound along with the vowel synthesis parameters. The functionality of the 8 buttons is:

1. a popupmenu button to select the vowel sound used for synthesis from among the 10 choices on the menu selection,
2. an editable button that specifies the sampling rate of the synthesized vowel sound, `fs`; (the default sampling rate is `fs=6000` samples per second),

3. an editable button that specifies the starting pitch frequency, `PF_Beg`, in Hertz; (the default value of `PF_Beg` is 100 Hz),
4. an editable button that specifies the ending pitch frequency, `PF_End`, in Hertz; (the default value of `PF_End` is 150 Hz),
5. an editable button that specifies the frame shift, R_m, in msec; (the default value of R_m is 10 msec);
6. a pushbutton to run the code and display the results of vowel synthesis on the eight graphics panel displays,
7. a pushbutton to listen to the synthetic vowel sound created by this program,
8. a pushbutton to close the GUI.

Vowel Synthesis – Scripted Run

A scripted run of the program 'Vowel_Synthesis_GUI25.m' is as follows:

1. run the program 'Vowel_Synthesis_GUI25.m' from the directory 'matlab_gui\vowel_synthesis',
2. using the popupmenu button, select the vowel sound `/IY/` for synthesis,
3. using the editable buttons, set the values for the synthesis sampling rate as `fs=10000`, for the pitch frequency starting value as `PF_Beg=100` Hz, for the pitch frequency ending value as `PF_End=150` Hz, and for the frame shift as $R_m = 10$ msec,
4. hit the 'Run Vowel Synthesis' button to compute and display the synthesis waveforms on the eight graphics panels,
5. hit the 'Play Synthetic Vowel' button to listen to the vowel sound created by this program,
6. experiment with different choices of vowel sound, and different values of the synthesis parameters, `fs`, `PF_Beg`, `PF_End` and R_m,
7. hit the 'Close GUI' button to terminate the run.

An example of the graphical output obtained from this exercise using the vowel resonances for the /IY/ vowel is shown in Figure 1. The graphics panels show the pitch period contour in the top left graphics panel; the glottal pulse in the middle left graphics panel; the convolved vowel sequence in the lower left graphics panel; the pitch excitation signal in the top right graphics panel; the vowel impulse response in the second right graphics panel; the vowel log magnitude frequency response in the third right graphics panel; the convolved vowel log magnitude frequency response in the fourth right graphics panel; and the convolved vowel linear magnitude frequency response in the bottom right graphics panel.

Vowel Synthesis – Issues for Experimentation

1. run the scripted program with a range of vowel sounds; is there any preference for one or more of the vowel sounds?
2. vary the pitch period contour (via the initial and final pitch editable buttons) to create rising, falling, and neutral pitch contours. Does the vowel sound more natural as the pitch varies over time? What characteristic of the sound is improved when the pitch period varies over a broad enough range?
3. change the sampling rate from 6000 Hz to 10000 Hz and then to 20000 Hz. What is the impact of changing sampling rate for this synthetic sound? Is there a highest frequency beyond which there is no perceptual difference as the sampling frequency is increased above this value?

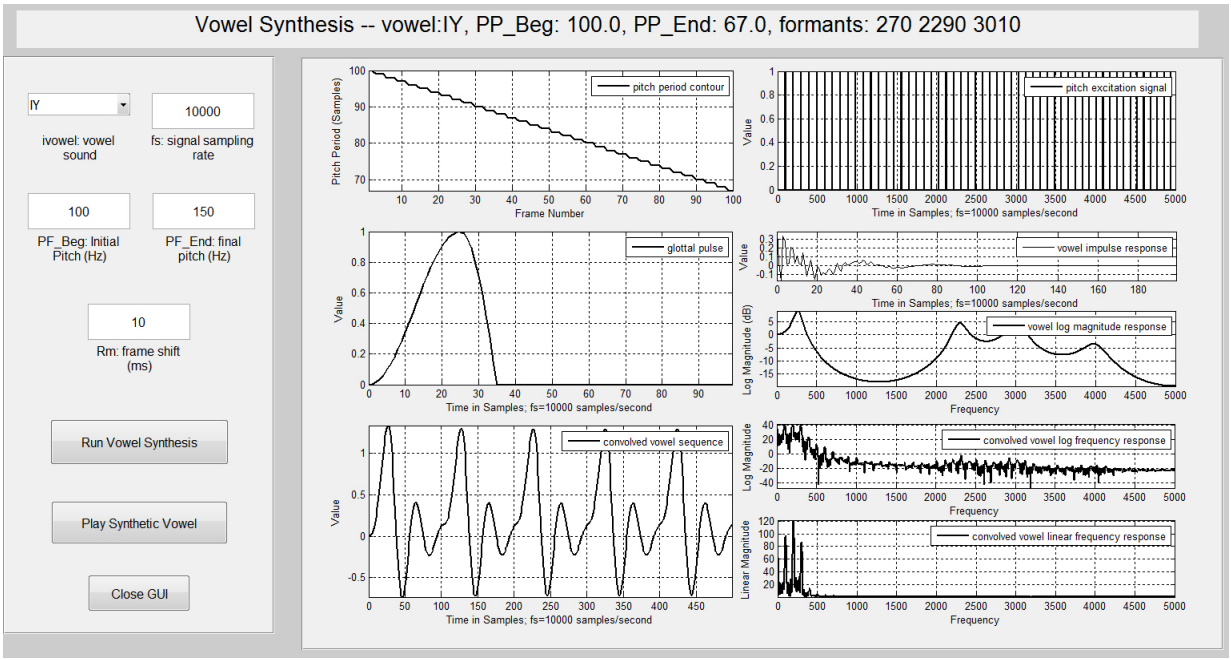

Figure 1: Graphical output from the synthesis of vowel sounds using the vowel sound /IY/ as the basic synthesis unit. The graphics panels show the pitch period contour in the top left graphics panel; the glottal pulse in the middle left graphics panel; the convolved vowel sequence in the lower left graphics panel; the pitch excitation signal in the top right graphics panel; the vowel impulse response in the second right graphics panel; the vowel log magnitude frequency response in the third right graphics panel; the convolved vowel log magnitude frequency response in the fourth right graphics panel; and the convolved vowel linear magnitude frequency response in the bottom right graphics panel.