MATLAB Exercise – Play Pitch Period Contour

Program Directory: `matlab_gui\play_pitch_contour`
Program Name: `play_pitch_GUI25.m`
GUI data file: `play_pitch.mat`
Callbacks file: `Callbacks_play_pitch_GUI25.m`

The goal of this MATLAB Exercise is to try to separate out the main properties (primarily pitch and intensity) of the speech excitation function (as estimated using LPC analysis) from the properties of the vocal tract (primarily the resonance structure associated with the spoken sound). To achieve this goal the proposed MATLAB program, using either a pre-recorded speech `.wav` file or a newly recorded file (done within this program), estimates the pitch period contour of the speech file and creates a pseudo-excitation function of pitch pulses (separated according to the pitch period estimate) and random noise. The program estimates the pitch period contour of the speech utterance and creates the two-state excitation sequence which is then convolved with the impulse response of a selected vowel sound (any of 10 standard vowels), and the resulting pitch modulated vowel is played out, thereby emphasizing the speech signal rhythm and tone without having the actually spoken sounds reproduced. In a sense, this program plays a speech melody (as obtained from the pitch period contour) without preserving the intelligibility of the spoken input utterance.

Play Pitch Period Contour – Theory of Operation

This exercise estimates the pitch period contour of either a pre-recorded speech utterance, or a newly recorded speech utterance, using a cepstral pitch detector of the type discussed in Chapter 10 of TADSP. Using the estimated pitch period contour, the exercise creates a two-state (voiced or non-voiced) excitation function that preserves the pitch period during voiced speech regions, and uses random noise during non-voiced regions. The two-state excitation function for the speech utterance is convolved with a fixed impulse response from one of the ten standard vowels of English (as represented by a set of four formant center frequencies and four formant bandwidths). The exercise plays out the resulting pitch-modulated vowel sound to get some sense as to the rhythm and tone of the spoken input utterance, without using any of the information about the vocal tract resonances.

Play Pitch Period Contour – GUI Design

The GUI for this exercise consists of three panels, 3 graphics panels, 1 title box and 13 buttons. The functionality of the three panels is:

1. one panel for the graphics display,
2. one panel for parameters related to the pitch period detection, and for running the program,
3. one panel for playing out the original speech signal and the vowel-modulated pitch contour.

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

1. the original speech waveform,
2. the median-smoothed pitch period contour for the spoken utterance,
3. the median-smoothed confidence score contour for the spoken utterance.

The title box displays the information about the selected file along with the set of short-time analysis parameters. The functionality of the 13 buttons is:

1. a pushbutton to select the directory with the speech file that is to be analyzed using short-time analysis methods; the default directory is 'speech_files',
2. a popupmenu button that allows the user to select the speech file for analysis,

3. an editable button that specifies the sampling rate for speech recording, `fsr`; (the default value is `fsr=6000` samples per second),

4. an editable button that specifies the number of seconds of recording, `nsec`; (the default value is `nsec=2` seconds of recording),

5. a pushbutton that initiates speech recording, after clicking the OK message from a text message box that appears on the screen,

6. an editable button that specifies the frame duration, L_m, (in msec) for short-time analysis; (the default value is $L_m = 40$ msec),

7. an editable button that specifies the frame shift, R_m, (in msec) for short-time analysis; (the default value is $R_m = 10$ msec),

8. a popupmenu button that specifies the pitch period range for the pitch detector being either male pitch range, female pitch range, or combined pitch range; (the default is male pitch range),

9. a popupmenu button that lets the user choose the vowel sound, `ivowel`, that is used to modulate the pitch period contour; (the default vowel sound is /IY/, but all 10 vowel sounds are available for synthesis),

10. a pushbutton to run the code and display the original speech signal, the median-smoothed pitch period contour, and the median-smoothed confidence score contour in the three graphics panel displays,

11. a pushbutton to play the original speech signal,

12. a pushbutton to play the pitch-modulated vowel signal,

13. a pushbutton to close the GUI.

Play Pitch Period Contour – Scripted Run

A scripted run of the program 'play_pitch_GUI25.m' is as follows:

1. run the program 'play_pitch_GUI25.m' from the directory 'matlab_gui\play_pitch_contour',

2. hit the pushbutton 'Directory'; this will initiate a system call to locate and display the filesystem for the directory 'speech_files',

3. using the popupmenu button, select the speech file for short-time feature analysis; choose the file 'test_16k.wav' for this example,

4. for recording a new speech file, set the initial values for the sampling rate to `fsr=10000` Hz, and for the number of seconds of recording to `nsec=2`; hit the 'Record Speech' button to initiate recording and click 'OK' on the text message box that appears on the screen in order to begin recording,

5. using the editable buttons, set the initial values for the short-time analysis parameters of frame length to $L_m = 40$ msec, and for frame shift to $R_m = 10$, msec;,

6. using the popupmenu button, set the pitch range to 'male pitch range',

7. using the popupmenu button, set the vowel sound for synthesis to /IY/,

8. hit the 'Run Play Pitch Contour' button to compute and display the original speech signal (top graphics panel), the median-smoothed pitch period contour (middle graphics panel), and the median-smoothed confidence score (bottom graphics panel), and to compute the pitch modulated vowel based on the formants of the designated vowel sound and the pitch contour of the designated speech signal, as estimated using a cepstral pitch detector,

9. experiment with different choices of speech file, with different vowel sounds, and with different values for L_m and R_m,
10. hit the 'Close GUI' button to terminate the run.

An example of the graphical output obtained from this exercise using the speech file 'test_16k.wav' is shown in Figure 1. The graphics panels show the speech waveform (top graphics panel), the median-smoothed pitch period contour as estimated using a cepstral pitch detection algorithm, and the median-smoothed confidence score contour, again estimated using a cepstral pitch detection algorithm.

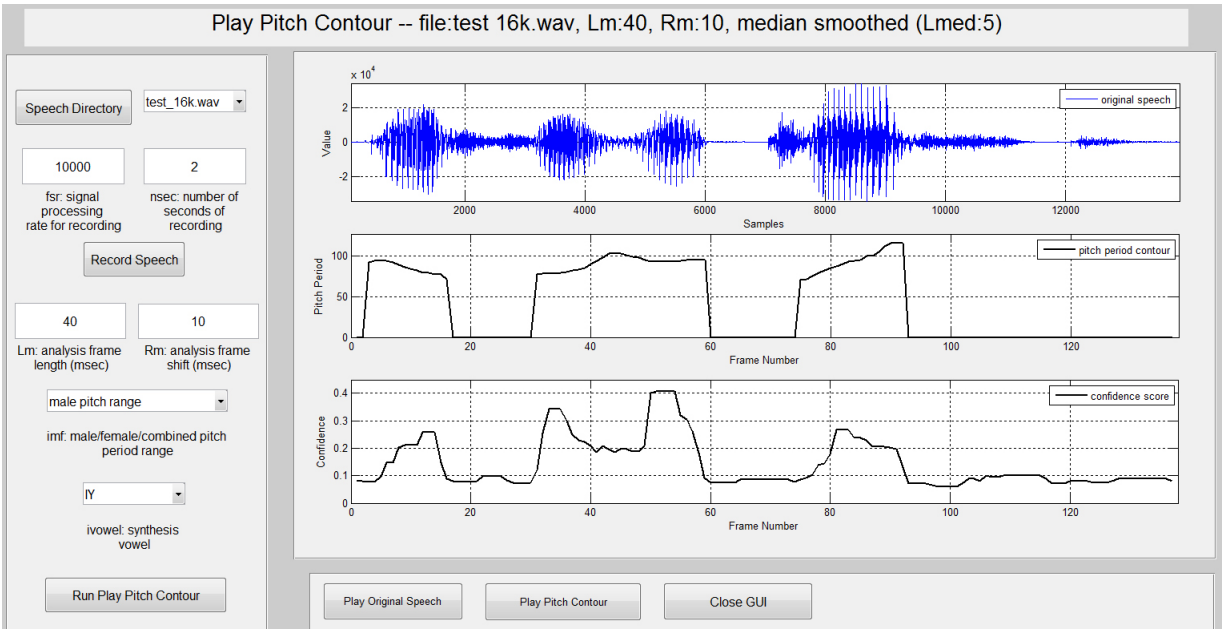

Figure 1: Graphical output from this exercise. The speech input file is 'test_16k.wav' and the vowel sound is the /IY/ vowel. The upper graphics panel shows the original speech waveform; the middle graphics panel shows the median-smoothed pitch period contour; and the bottom graphics panel shows the median-smoothed confidence score for the pitch period estimates.

Play Pitch Period Contour – Issues for Experimentation

1. Select a speech signal such as 's5.wav'. Listen to the sound, paying attention to pitch changes throughout the utterance.
2. Run the exercise on this signal. Observe the detected pitch contour. How well did the pitch detector work on this utterance? Now listen to the result of using the pitch contour as the excitation and an /IY/ vowel as the speech system. Can you hear the pitch changes in this vowel sound? Listen to the original signal again. Are the pitch changes the same as for the vowel?
3. Try using the same pitch contour with other vowels and repeat the above analyses.
4. Record a signal by "singing" a vowel sound such as /AA/. While singing, try to make the pitch variations as large as you can. Now apply these variations in pitch to a synthetic /AA/ vowel. Again compare the sound to that of the original sentence.
5. Repeat the above exercises with yet another synthetic vowel sound.