

## **CSC784M – Machine Problem 2 (Due 5-Mar-2019, 6:00 PM)**

**INSTRUCTION:** *This activity may be done by pair or individually. You are only allowed to use MATLAB or Octave. For this machine problem, you will need to download the speech signal 'speech\_MP2.wav' /Tutugtog ba si Francisco?/. Documentation for this MP should contain: M-file scripts or functions, results, screenshots, discussions, and answers to questions. Submit your documentation (in \*.pdf file format) and M-file(s) online thru Canvas before the deadline. Submit a printed copy of your documentation at the start of the class on 5-Mar-2019.*

1.) For the given speech data ('speech\_MP2.wav'), perform segmentation such that each segment or frame has a 50-millisecond duration, and that every two successive frames have an overlap of 40 milliseconds.

a.) Write an M-file that examines each frame whether it contains a voiced or an unvoiced sound. Your program should leave the signal unaltered for frames that were classified as voiced sound, otherwise mute (i.e., set the signal to zero) the frames that were classified as unvoiced sound. Then for each frame that contains voiced sound, your program should estimate the pitch (F0). The program should finally make subplots of the resulting signal (i.e., the one with voiced sounds only) and the pitch contour. Discuss your pitch estimation algorithm. Why did you choose this algorithm? (Justify your answer.) Is the pitch (F0) generally related to vowel sound class, whether it is /a/, /i/, /o/ or /u/? (Show/discuss your proof.)

b.) Use PRAAT to estimate the pitch contour from the given speech data. (Simply use the default or the recommended settings for pitch estimation in PRAAT.) How does your result in part (1.a) compare with PRAAT's result? Discuss any similarities or differences, and present your analysis.

2.) Use the same speech data and segmentation as in part (1)

a.) Write an M-file that estimates the F1 and F2 formant frequencies for each frame with voiced sounds. The program should make subplots of the given speech signal and the F1 and F2 contours. Discuss your F1 and F2 estimation algorithm. Why did you choose the algorithm? (Justify your answer.) Are the F1 and F2 generally related to vowel sound class, whether it is /a/, /i/, /o/ or /u/? (Show/discuss your proof.)

b.) Use PRAAT to estimate the F1 and F2 contours from the given speech data. (Simply use the default or the recommended settings for formant estimation in PRAAT.) How does your result in part (2.a) compare with PRAAT's result? Discuss any similarities or differences, and present your analysis.