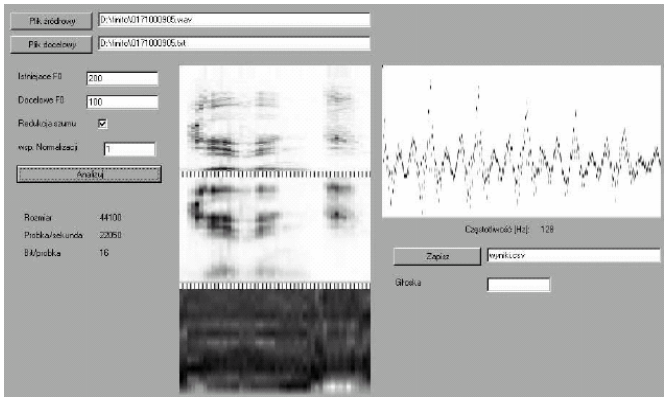
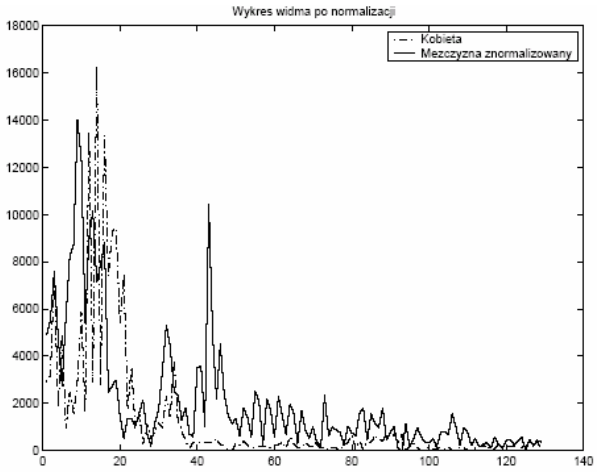
Project no. 10 (2 persons)

Detection of speaker – dependent features





Following steps should be performed:

(a) (P1) Read and write a .wav file, acquire different learning samples of speech for different speakers – write them as .wav files. Perform initial segmentation into phonemes and silence. Make detection of voiced and unvoiced phonemes in the time signal. Present the waveform and the 2-D spectrogram in graphic windows.

(b) (P2) Perform the windowed Fast Fourier Transform (FFT). Present the spectrogram image. Detect the maxima for every window – detect the formants - present the formants in the spectrogram image.

(b) (P1, P2) Propose and test a normalization procedure for the MFCC feature-detection procedure on base of detected border positions of formants. Prepare a final report.