

Audio data of 1 sec (44100 samples) is dumped into input FIFO.

Append latest 1 sec data with previous 0.98 sec data.

Compute super feature vector from MFCC features extracted for audio segment of 1.98 sec.

Compute R^{12} MFCC feature vector for each 1024 point FFT, computed on 1.98 sec of audio clip. This results in 169×12 matrix.

