ENG4053

Assignment 1

Digital Signal Processing

Fourier Transform

University of Glasgow, School of Engineering

Bernd Porr & Nick Bailey

This assignment is about the Fourier Transform, how to enhance your voice and reduce artefacts.

Form groups of two, work together and submit one report. Enter your team into the Wiki on moodle listing your names and matriculation numbers. Use the moodle forum to find team mates.

Record the voice with a handheld mic of one member of your group as uncompressed WAV.
Record the sentence twice:
a) Mic / speaker distance: 1m
b) Mic / speaker distance: 2cm
and say one complete sentence. It can be any language. Make sure that you record at least at 44kHz or at a higher sampling rate and that the audio is not clipping. The full audio spectrum up to 20kHz needs to be available. Low quality MP3 downloads from websites converted to WAV are not allowed and won't be marked. Reports based on recordings at sampling rates below 44kHz or downloads will not be marked.

1. Load the audio samples into python.
    1. plot the audio signals in the time domain (linear axis: normalised amplitudes -1..+1 vs time) and
    2. in the frequency domain (logarithmic axis for both frequency and amplitude in dB) with proper axis labels.

2. Use a *vector*-drawing program (Inkscape, Illustrator, drawio, ...) and mark
    1. the peaks in the spectrum which correspond to the *fundamental* frequencies of the vowels spoken.
    2. Mark up the frequency range which mainly contains the *consonants* up to the highest frequencies containing them.
    3. Mark up the spectral frequencies which most likely just contain noise.

    Provide brief explanations. **30%**

3. Use the speech audio from the two recordings and improve the quality of the voice for both of these recordings. Do this by boosting certain frequency bands and/or damping others. Provide an explanation. **30%**

4. Use the speech audio from the two recordings and remove noise from both of these recordings. Do this by removing certain frequency bands. Provide an explanation. **30%**.
5. Compare the problems and advantages of having a microphone close by and far away in terms of post processing. **10%**.

The report should be brief, concentrating on the technical aspects and why you have done the different steps. Do not add generic theory about voice or Fourier Transform. Just describe the method and the result. Complete PYTHON code *must* be included in the appendix and submitted via moodle alongside the report. All figures inline in the report must be high quality graphics in vector format. Blurry jpeg figures or screenshots will not be marked. Figure out early how to add high quality graphics to your document such as EPS or SVG. Submission must be PDF.

Upload your code, data/WAV files to moodle in form as a single zip file. Follow *exactly* the naming conventions for all files as specified on moodle. The scripts will be tested under Linux from the commandline (so not Spyder, Pycharm or VScode). Make sure your code is platform independent and does not contain absolute paths. Code that crashes will result in low marks. The same applies for code which won't display any plots, for example forgetting "plt.show()". Your audio files must be original ones and WAV 16 bit. No high level python signal processing / filtering commands are allowed except of the numpy FFT and IFFT commands.

Deadline: 16th Oct 2023, 3pm on moodle.