ENG5027- Digital Signal Processing

Assignment 1: Fourier Transform

This assignment is about the Fourier Transform.

Teams

In this assignment you are expected to form groups of 3 or 4 people and work together to submit a single report for this assignment. To form a team you should enter your name and matriculation number into the wiki Moodle listing then use this Moodle wiki to find teammates.

Audio Specification

In your groups, record the voice of one member of the as an uncompressed WAV file.

Recording specification:

- Recording must be based on the voice of one member of the team
- Recoding should be of one complete sentence in any language
- Recording must be uncompressed WAV file
- Recording must be of a sampling Rate of 44kHz or higher
- Recording must be audio must not exhibit clipping
- Recording must contain the full audio spectrum up to 20kHz

If reports are generated using files that do not conform to the recording specifications, they will not be marked. For example,

- Reports based on low quality MP3 files.
- Reports based on sampling rates below 44kHz.
- Reports based on downloaded files from the internet.

Assignment Tasks

Task 1: Loading Audio into Python

1. Read the audio samples into a python application
2. Plot the audio signal
    a. Plot 1: normalised amplitudes vs time using a linear axis in the time domain
    b. Plot 2: amplitude (dB) vs frequency using logarithmic axis in the frequency domain

NOTE: all plots should have proper and appropriate labels.

This section is worth: **5%**

Task 2: Audio Analysis

Use a drawing program (Inkscape, Illustrator, drawio, …) to mark up the plots generated in task 1. Specifically,

1. Mark the peaks in the spectrum which correspond to the fundamental frequencies of any spoken vowels present in the sample
2. Mark the frequency range which mainly contains the consonants up to the highest frequencies containing them
3. Mark the whole speech spectrum containing the vowels, consonants harmonics.

Additionally, please provide brief explanations for each marked plot.

This section is worth: **15%**

Task 3: Fourier Transform

In this task, you will use the Fourier transform to modify the audio signals with the goal of increasing the voice quality. Using the audio from the previous tasks you must determine the region of the highest harmonic voice frequencies in the spectrum and increase the amplitudes of those frequencies with the help of the Fourier Transform.

This section is worth: **30%**

Task 4: Vowel Detector

In this task you will write a simple python program to detect at least two different vowels in the voice recording. Specifically,

- Write a function which takes an audio file as input and outputs the corresponding vowel as a string.
- This function should be called twice or more with two different WAV files
- This function must be defined as:

```
# function to identify vowels
 def voweldetector(wavfile):
        #<code here>
```

This section is worth: **50%**

Submission

Report Format

- The report should be brief and concentrate on the technical aspects and why you have performed the different steps: focus on the method performed and the result.
    - Do not add generic theory about the Fourier Transform or voice.
- Complete PYTHON code must be included as an appendix and submitted with the report. Reports without PYTHON code will not be marked.
- All figures in the report must be high quality graphics in vector format. Blurry jpeg figures or screenshots will not be marked.

- Report submissions must be in pdf format

Python Code

- Submitted code must be platform independent and not contain absolute paths.
    - Code that crashes will receive reduced marks
    - Code that does not display plots will receive reduced marks
- The only allowed PYTHON signal process functions are FFT and IFFT
    - Other signal processing / filtering functions are not allowed

Zip File Naming & Contents

One submission per *team*. Every team submits only one zip file. Filename convention: use exactly these filenames. Any other file will be ignored.

- report.pdf
- original.wav (16bit PCM)
- voice_enhancer.py
- improved.wav (16 bit PCM)
- voweldetector.py
- vowel1.wav (16 bit PCM)
- vowel2.wav (16 bit PCM)

Feel free to split your project into more files, modules etc but only the python scripts above will be tested if they will work. All files need to be zipped into a single zip with your matric numbers in it:

fft_12234567a_12345886b_9957754c.zip.

Make sure that the python files run directly from the command line, for example by typing "python voice_enhancer.py". -- And remember the submission must be original work.

Submissions in the wrong format will not be marked.

Submission Deadline

The submission deadline is Tuesday 18th October 2022 @ 3pm. Submissions must be made via Moodle.