Audio and Video Signals (AVS): *Module 1 – Audio Signals (AS)*	*Written test of February 6, 2023*
Name: **ID (Personal Code):**	

Please answer **thoughtfully**, **pertinently** and **briefly.**

Q1. [8] A customer asks you to design an audio-conferencing system that implements two features:

- Echo cancellation to improve the audio quality within the room.
- Source localization to make it possible to automatically point a camera to the speaker in case of video recording.

To implement echo cancellation, you decide to exploit a classic approach based on Wiener filtering.

a) [1] Sketch the block diagram of a classic echo cancellation system based on Wiener filter.
b) [3] Compute the optimum filter coefficients assuming the following:
   - Room impulse response of the echo source $H(z) = 0.6 + 0.2z^{-1}$;
   - Target speaker signal $s(n)$ and noise signal $u(n)$ are real-valued uncorrelated zero-mean white noises with variances $\sigma_s^2 = 0.8,\ \sigma_u^2 = 0.4$.

To implement the source localization system, you decide to use a uniform linear array (ULA) composed by 130 microphones.

c) [1.5] Which is the distance between two consecutive microphones to avoid aliasing at 3400 Hz?
d) [1] How long is the entire array?
e) [1.5] Which is the angular resolution of the array at 340 Hz?

Q2. [4] Consider the problem of synthesizing audio signals by means of different approaches.

a) Describe how Wavetable Synthesis method works.

b) Report possible pros and cons of Wavetable Synthesis.

c) Describe how Granular Synthesis works.

Q3. [4] Consider the problem of classifying a set of sounds into two different classes (e.g., music vs. speech). To solve this problem, you are given a set of labeled audio recordings, and you plan to use a supervised machine learning approach.

a) Sketch the block diagram of a possible pipeline that can be used to solve the problem. Consider both the training and test pipeline.

b) Report two examples of features that can be extracted from audio signals.

Q4. [7] Each answer can be either TRUE or FALSE. Incorrect answers give you -1 points, unanswered 0 points, correct answers +1 point. Provide a brief optional comment when in doubt.

a) [T] [F] A uniform linear array (ULA) composed by three microphones spaced 5 cm apart one from the other can possibly measure a TDOA between two microphones of 0 (zero) seconds.

b) [T] [F] A L=2 factor window can be used to resolve spectral peaks at 100 Hz distance in the frequency domain if the frame under analysis is 320 sample long and has been acquired at 8000 Hz sampling frequency.

c) [T] [F] Consider a sinusoid with a frequency of 440 Hz. By applying time-scaling to this sinusoid, we obtain a sinusoid with the same frequency.

d) [T] [F] Consider the problem of background noise removal by means of Wiener filtering. The Wiener filter tends to attenuate spectral bands in which the power spectrum of the target signal is significantly higher than that of the noise.

e) [T] [F] It is possible to convolve a room impulse response (RIR) with a dry audio signal to introduce reverberation effects.

f) [T] [F] Dattoro industry standard scheme can introduce reverberation effects on audio signals.

g) [T] [F] Given a room and a microphone, two different sources may lead to two different room impulse responses (RIRs).

Q5. MATLAB [8] Considering a sampling frequency of 4 kHz, implement the following steps.

Define some signals

- [0.5] Generate a sinusoidal signal `y_1` with the following characteristics: frequency 440 Hz, amplitude 1, phase 0, duration 1 second.
- [0.5] Generate a sinusoidal signal `y_2` with the following characteristics: frequency 880 Hz, amplitude 0.5, phase 0, duration 0.5 second.

Plot

- [1] Plot both signals in the time domain expressing the time axis in seconds (hint: you can use two separate figures).
- [1] Plot the spectrum of each signal expressing the frequency axis in Hz (hint: you can use two separate figures).

Windowing

- [0.5] Define a 20 ms Hamming window.
- [0.5] Extract the first frame of 20 ms from `y_1` and window it.
- [1] Considering an overlap of 10 ms, compute how many frames you can extract from both `y_1` and `y_2`.

Feature computation

- [1] Compute the zero-crossing rate (ZCR) of both `y_1` and `y_2`.
- [0.5] Comment the computed (or the expected) ZCR in relation to the signals' length and frequency.

Filtering

- [0.5] define the two-sample filter [-0.5, 0.5].
- [1] Apply the linear filter to `y_1` using any valid technique of your choice.