Audio and Video Signals (AVS): ***Module 1 – Audio Signals (AS)***	*Written test of January 12, 2023*
Name: **ID (Personal Code):**	

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

Q1. [8] Consider a teleconferencing system comprising a stereo loudspeaker system a proximity microphone capturing the talker speech s(n) and an environmental microphone. The goal of the acoustic echo cancellation is to extract from the environmental microphone signal y(n), the audience speech u(n) and noise n(n) removing the contribution of the signal x(n) emitted by the two loudspeakers and that of the talker.

a) Sketch the block diagram of the echo cancellation system and briefly comment it.

b) Assume that the response of the loudspeakers is known

$$h_1(n) = 0.8x(\mathrm{n}) + 0.4x(n-1)$$

$$h_2(n) = 0.2x(\mathrm{n}) + 0.1x(n-1)$$

In addition, the talker signal acquired by the environmental microphone is modelled as

$$s'(n) = s(\mathrm{n}) + 2s(n-1)$$

The signals are defined for simplicity as white gaussian noises with

$x(n) \sim \mathcal{N}(0, \sigma_s^2 = 0.1) \in \mathbb{R}$

$u(n) \sim \mathcal{N}(0, \sigma_s^2 = 0.5) \in \mathbb{R}$

$s(n) \sim \mathcal{N}(0, \sigma_s^2 = 1) \in \mathbb{R}$

$n(n) \sim \mathcal{N}(0, \sigma_s^2 = 0.2) \in \mathbb{R}$

b1) Compute the optimum filters with $M = 2$ taps to estimate the audience and noise contribution removing the stereo loudspeaker contribution.

b2) Compute the optimum filters with $M = 2$ taps to estimate the audience and noise contribution removing the talker contribution.

Q2. [4] Consider the problem of synthesizing a reverberated audio signal.

a) Define the concept of Room Impulse Response (RIR), sketching one as an example.

b) Define the concept of reverberation time T60.

c) Explain how it is possible to use a RIR to apply reverberation to a dry audio track.

Q3. [4] Consider the problem of applying time and pitch scaling to an audio signal.

a) Define the concept of time scaling.

b) Define the concept of pitch scaling.

c) Considering a speech signal, describe the pitch synchronous overlap and add (PSOLA) procedure explaining if it applies time or pitch scaling and why.

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] Consider the problem of resolving spectral peaks in an audio signal composed by a sum of sinusoids. The length in samples of the signal under analysis impacts on the frequency resolution.

b) [T] [F] Consider the use of linear predictive coding (LPC) to estimate the spectral characteristics of a signal. By increasing the LPC order, the spectral estimates become smoother resembling more and more the spectral envelope.

c) [T] [F] Consider sound synthesis by means of subtractive synthesis (source-filter model). The input excitation signal (source) is typically a sinusoid.

d) [T] [F] A uniform linear array composed by 10 microphones spaced 5 cm apart one from the other does not suffer from aliasing if used to localize sources that emit signals with a maximum frequency of 4000 Hz.

e) [T] [F] In order to use a parametric method for direction of arrival (DOA) estimation, the number of sources must be known.

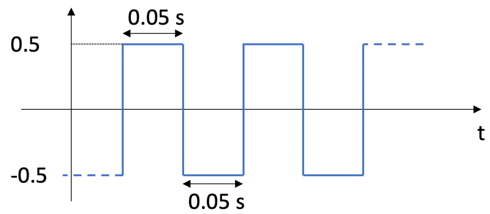
f) [T] [F] In a time-variant fractional delay line it is more desirable to use a linear interpolator than an all-pass interpolator if the delay changes randomly rather than being constant.

g) [T] [F] Consider psychoacoustic effects. Masking between tones only happens if the tones are played simultaneously.

Q5. MATLAB [8]

Define some signals:

- [1] Consider a sampling frequency of 4 kHz, generate a 0.5-second-long cosine with frequency 500 Hz and amplitude equal to 0.7. This is your signal `y_1`.
- [1.5] Generate the signal `y_2` as a 1-second-long rectangular wave with the characteristics shown in the figure (i.e., amplitude ranging from -0.5 to 0.5, period of 0.1 s, and duty cycle of 50%).



- [0.5] Obtain the signal `y` as the sum of `y_1` and `y_2`. Decide how to deal with the difference in length.
- [1] Compute the zero-crossing rate of `y`
- [0.5] Extract one frame `y_w` from `y` using a 200 ms Hamming window.
- [0.5] Plot `y_w` in the time domain correctly expressing the time axis in seconds.
- [1] Plot the spectrum of `y_w` correctly expressing the frequency axis in Hertz.

Apply a filtering operation in the frequency domain on `y_w`:

- [1.5] Obtain the signal `y_filt` that consists of `y_w` filtered with the filter [-0.5, 1, -0.5].
  - The filtering operation must be implemented in the frequency domain.
  - Work with an amount of samples that is a power of two in the frequency domain.
- [0.5] Can you tell if the filter is a high-pass or low-pass filter? How?