TÉCNICO LISBOA

Learning-based Multimedia Processing

2022/2023

IST number: _____

Name: _____

Quiz #1 **Duration: 20 minutes**

Provide clear, legible, and succinct answers.
Always justify your assumptions.

Questions

1. For telephonic communications an acoustic speech signal can be converted into an analogue electrical signal using a microphone. The resulting analogue speech signal to be transmitted can be limited to the frequency range from 20 Hz to 4000Hz.

a) What needs to be done so that this speech signal can be transmitted using a digital communications system? (specify the main operations involved)

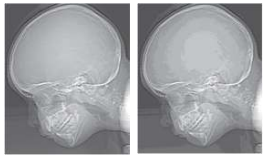
b) What is the typical bitrate (bit/s) resulting from the conversion from analogue to digital speech? (justify all the values you consider)

2. The human visual system captures the light coming from a scene which reaches the retina.

a) What are the two types of light "sensor" cells present in the retina, and what are the main characteristics of each one?

b) Are there areas in the retina that are favoured in terms of human vision? Or areas of the retina that do not contribute to human vision? (briefly explain why)

3. Consider an original scene (left) and an acquired image (right). What can explain the acquired image characteristics?
How should the acquisition system be modified to solve the problem?



4. Consider a camera with a small CCD sensor, with a resolution of 200x200 *pixels*, where each pixel occupies an area of 5μm x 5μm. The camera is used to take a picture of a 10cm x 10cm object, using a lens with focal length of 1 cm. How far away from the object should the camera be placed so that the object image completely occupies the camera sensor area?
(Start by drawing a figure illustrating the image formation process; 1 μm = 1×10^{-6} *m*)

5. Consider a digital image of size 1000 x 1000 pixels, each pixel represented using 8 bit/pixel. Would it be possible to store this image using at most 80 kbit? What type of operations can be considered for this purpose (if any)?

Solution:

1 –

a) sampling + quantization;

b) 4000 x 2 samples/s x 8 bit/sample = 64 kbit/s)

2 –

a) rods and cones;

b) fovea: more detail (concentration of cones); blind spot in the optical nerve location

3 –

intensity resolution is too low – increase the number of levels used to represent the intensity of each pixel; i.e., increase the number of bit/pixel

4 –

$h/f=H/d \iff d = f \times H/h \iff d = 0.01 \times 0.1/(200 \times 5 \times 10^{-6}) = 1\ m$

5 –

The raw image occupies 1000 x 1000 x8 bit = 8 Mbit;

The image can be compressed, exploiting irrelevancy and redundancy