MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO, Alameda

Academic Year 2023/2024 – 1st Semester, Responsible: Prof. Fernando Pereira

1st Exam – 17th January 2024 (Wednesday), 6pm

The marks should be out before **19th January (Friday), 12pm** at the CMul Web page. The exam scoring checking session will be on **19th January (Friday) at 4pm**, room LT4.

The exam is **90 minutes long**. Answer all the questions in a detailed way, **including all the computations performed and justifying well your answers**.

Don't get 'trapped' by any question; move forward to another question and return later. ***Boa sorte !***

I (1.0 + 1.0 + 1.0 + 1.0 + 1.0 + 2.0 val. = 7.0 val.)

Consider a videotelephony communication using Recommendation ITU-T H.261. The video sequence is coded with a CIF spatial resolution and a frame rate of 10 Hz at a (constant) channel bitrate of 128 kbit/s. The bits for each frame are uniformly generated in the time interval that the encoder usually dedicates to encode each image. At the encoder, the bits wait for transmission in an output buffer.

Answer the following INDEPENDENT questions …

a) Assuming that the first frame produces 16000 bits and the buffer size is 20000 bits, what is the **maximum number of bits** that the second frame may produce ? (R: 29600 bit)

b) Assuming that the first frame produces 28000 bits and the buffer size is 20000 bits, what is the **maximum number of bits** that the third frame may produce ? (R: 30400 bit)

c) Assuming that the first frame produces 18000 bits, what is the **minimum size of the buffer** ? (R: 5200 bit)

d) Assuming that the buffer size is 25600, what is **latest tim**e the full set of bits for the second frame may be received at the decoder ? (R: 400 ms)

e) Why are the **motion vectors differentially coded** ? State one positive and one negative impact of this decision. (R: increase compression efficiency and reduces error resilience)

f) How and why would the **video quality change** (notably the average quality) if: i) the channel rate increases; ii) the frame rate decreases; iii) the spatial resolution becomes QCIF; iv) the output buffer size increases.

II (2.0 + 2.0 + 1.0 + 1.0 = 6.0 val.)

Consider that your company is contacted to design a videoconference system between the various main locations of a bank. The spatial resolution is CIF (352×288 luminance samples), 4:2:0, at 12.5 Hz, with the usual number of bits per sample. Assume that you have available, offering the target video quality, two solutions:

1. H.261 based solution with average compression factors of 25 and 35 for the luminance and chrominance, respectively; the critical compression factors (for the images spending more bits) are 20 and 25 for the luminance and chrominances, respectively.

2. MPEG-2 Video based solution with N = M = 3 with average compression factors of 25 and 35 for the luminance and chrominance, respectively, for the I frames, and 30 and 45 for the luminance and chrominances, respectively, for the P and B frames. The critical compression factors are 75% of the average compression factors.

Assume that the transmission rate is always the same as the coding rate.

a) Determine the **bitrate and acquisition-visualization delay** for the H.261 based solution. (R: 550,326 kbit/s and 103 ms)

b) Determine the **bitrate and acquisition-visualization delay** for the MPEG-2 Video based solution. (R: 483,816 kbit/s and 400 ms)

c) Assuming that your client always pretends to minimize the transmission rate, **what solution** from above would you select *as a function of the minimum acquisition-visualization delay value defined by the client*?

d) State and explain two key consequences (one positive and one negative) of changing N=M=3 (as above) to N=9, M=3 after describing the frame prediction structure.

III (1.0 + 1.0 + 1.0 + 1.0 + 1.0 + 1.0 + 1.0 = 7.0 val.)

Consider the audio codec specified in MPEG-1 Audio.

a) What is a **reasonable rate** required to code a stereo signal with 44 kHz sampling rate per channel? (R: 117,33 kbit/s)

b) Why is the MP3 coding architecture called a **hybrid coding architecture** ?

c) What is the **temporal post-masking effect ?** What is its typical duration ?

d) What does it happen if the 'spatial integrity' is lost when coding a stereo audio signal ?

e) Indicate and explain **two coding effects** that may happen if the audio encoder includes a psychoacoustic model which is conservative in the sense of exaggeratedly lowering the hearing thresholds associated to the audio masking effects ?

f) Considering the varied composition of a symphonic orchestra, **how would change the subjective quality** assessment associated to the music if the full bandwidth is kept while successively reducing the initial sampling rate from a value which starts being 2.5 times the full bandwidth and ends being the same as the full bandwidth?

g) What would happen in terms of **subjective quality assessment** if the full bandwidth is now reduced in the same proportion as the sampling rate starting from the same initial values as in f) ?