MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2020/2021 – 2nd Semester, Responsible: Prof. Fernando Pereira

2nd Exam – 8th July 2021 (Thursday), 8am

The marks should be out before **10th July (Saturday), 6pm** at the CMul Web page. If you want to check the exam scoring, please contact the course responsible by email **before 12th July (Tuesday)** to schedule a conversation.

The exam is **2.5 hours 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 val. = 5.0 val.)

Consider the JPEG standard to code photographic images with a 288×352 luminance resolution and 8 bit/sample.

a) Determine **the 'price' in bits per pixel** if 4:4:4, 4:2:2 and 4:2:0 color subsampling is used and **no compression is used**. (R: 24, 16, 12 bpp)

b) Determine **the average 'price' in bits per pixel** if 4:4:4, 4:2:2 and 4:2:0 color subsampling is used and compression is applied with a luminance compression factor of 20 and a chrominances compression factor of 25. (R: 1.04, 0.72, 0.56 bpp)

c) Determine **the average 'price' in bits for one single luminance sample** if compression is applied with a luminance compression factor of 20. Determine the same price for one single chrominance sample if compression is applied with a chrominance compression factor of 25. (R: 0.4, 0.32 bit per sample)

d) Determine the **total number of bits** that have to be spent to code all the components of a full 4:2:0 image if, on average, 4 DCT coefficients are coded per block and each DCT coefficient costs, on average, 5 bits; additionally consider that the EOB (End of Block) word costs 4 bits. (R: 57024 bit)

e) Determine the **total number of bits** that have to be spent to code all the components of a full 4:4:4 image with a single uniform/flat colour if, on average, each DCT coefficient costs 5 bits; additionally consider that the EOB (End of Block) word costs 4 bits. Explain your reasoning. (R: 42768 bit)

II (0.5 + 0.5 + 0.5 + 1.0 + 0.5 + 1.0 val. = 4.0 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 64 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 buffer size is 18000 bits, what is the **maximum number of bits** that the first frame may produce ? (R: 24400 bit)

b) Assuming that the first frame produces 15000 bits, what is the **minimum size of the buffer** ? (R: 8600 bit)

c) Assuming that the first frame produces 15000 bits and the buffer size is 18000 bits, what is the **maximum number of bits** that the second frame may produce ? (R: 15800 bit)

d) Assuming that the first frame produces 20000 bits and the buffer size is 15000 bits, what is the **maximum number of bits** that the third frame may produce ? (R: 14200 bit)

e) Assuming that the buffer size is 12800, what is **maximum number of bits** that frame 467 may produce ? (R: 19200 bit)

f) Assuming that the buffer size is 12800, what is **latest tim**e the full set of bits for the second frame may be received at the decoder ? (R: 400 ms)

III (1.0 + 2.0 + 1.0 = 4.0 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: 550.3 kbit/s; 103 ms)

b) Determine the **bitrate and acquisition-visualization delay** for the MPEG-2 Video based solution. (R: 483.8 kbit/s; 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*?

IV (0.5 + 1.0 + 0.5 + 1.0 + 0.5 + 0.5 = 4.0 val.)

Consider audio signals.

a) Indicate **two audio characteristics** which are numerically larger than the corresponding speech characteristics. (R: bandwidth and dynamic range)

b) What is the **main audio characteristic** justifying to use a higher bit depth (i.e. number of bits per sample) than for speech ? Explain the process involved. (R: larger dynamic range)

c) Explain in what consists the **temporal post-masking effect.**

d) What is the **main motivation** to increase the size of the audio frames in MPEG-1 Audio Layer 2 regarding Layer 1? Which performance metric suffers if we keep increasing the audio frames size ?

e) Would you prefer listen to 22 kHz bandwidth with appropriate sampling and 16 bit/sample audio OR 44 kHz sampling frequency and 12 bit/sample audio ? Why ? (R: first solution)

f) Considering the varied composition of a symphonic orchestra, how would change the subjective quality assessment associated to the music if the full bandwidth (22 kHz) is kept while successively reducing the sampling rate from a value which starts being 2.5 times the full bandwidth and ends being 1.5 times the full bandwidth? (R: reducing quality due to aliasing)

V (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 = 3.0 val.)

Consider the H.264/AVC video coding standard.

a) What is the offered **compression factor** if this codec spends around 2Mbit/s to code video with a spatial resolution of 576×720, 4:2:0 at 25 Hz ? (R: 62.2)

b) Explain what is a **practical implication** of using the Flexible Macroblock Ordering tool.

c) Why are motion vectors differentially coded but not across slices ?

d) Explain why Instantaneous Decoding Refresh (IDR) pictures and not just Intra pictures are essential to allow zapping in a broadcasting scenario.

e) For what reason was the Constrained Intra Coding Mode defined ?
f) Explain why is the H.264/AVC transform process called a hierarchical transform ?