TÉCNICO LISBOA

MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO, Alameda

Academic Year 2021/2022 – 1st Semester, Responsible: Prof. Fernando Pereira

2nd Exam – 24th February 2022 (Thursday), 6pm

The marks should be out before **25th February (Friday), 8pm** at the CMul Web page. The exam scoring checking session will be on **28th February (Monday) at 2.30pm**, room LT4.

The exam is **2 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 1152×1440 luminance resolution, 4:2:0 color subsampling and 8 bit/sample.

a) How many pixels, samples and blocks exist in this type of image. (R: 1658880; 2488320; 38880)

b) Determine the average price (in bits) of a pixel if a JPEG codec with a luminance compression factor of 25 and a chrominance compression factor of 15 is used; consider 12 bit/sample. (R: 0.88 bit/pixel)

c) Determine the average price in bits for a luminance block if a codec with a luminance compression factor of 25 and a chrominance compression factor of 15 is used. What would be the price variation for a chrominance block ? (R: 20.48; 13.66)

d) Determine the price in bits to code a gray scale version of a full image with only 64 levels of gray if the same compression factors as in c) are used. (R: 398131.2)

e) Determine the total number of bits that have to be spent to code only the chrominances components of one image if an average number of 4 DCT coefficients are coded per block and each coefficient costs, on average, 5 bits; additionally consider that the EOB (End of Block) word costs 4 bits. How many more bits would be needed to also code the luminance component if an average number of 5 DCT coefficients are coded per block and each coefficient costs, on average, 3 bits (and the same EOB is used). (R: 311040; 492480)

II (1.0 + 1.0 + 1.5 + 1.0 + 0.5 = 5.0 val.)

Assume that you are contacted by a company to design a digital storage system for video clips. The company requires some editing flexibility and needs to store the largest number of 3 minutes clips in a disk. The maximum access speed to the disk is 10 0 Mbit/s. The clips have 4K resolution with the following characteristics: 3840 × 2160 (Y), 4:2:0, 10 bit/sample at 25 Hz.

a) Assuming that you have at your disposal, providing the required video quality, a JPEG coding solution with average compression factors of 35 and 30 for the luminance and chrominances, respectively, determine the maximum access time for an image knowing that the compression factors for critical frames are 10% lower than average. (R: 41.69 ms)

b) Assuming now that you have at your disposal, providing the required video quality, a MPEG-2 Video coding solution with N=12 and M=4 with the following average compression factors:
   - I frames: 30 and 35 for the luminance and chrominances, respectively
   - P frames: 40 and 50 for the luminance and chrominances, respectively
   - B frames: 50 and 60 for the luminance and chrominances, respectively

   Determine the maximum access time for an image knowing that the compression factors for critical frames are 25% lower than average. (R: 214 ms)

c) Determine, justifying, which coding solution would you propose to your client if the target is only to maximize the number of clips stored in the disk. (R: MPEG-2 Video)

d) Assuming that your client always pretends to minimize the coding rate, what solution from above would you select *as a function of the maximum random access time requirement defined by the client*?

e) If the MPEG-2 Video coding solution changes to N=20 and M=4 (and not N=12 and M=4 anymore), explain without any computations if this change would increase or reduce its chances to be adopted ? (R: Reduce)

III (1.0 + 1.0 + 1.0 + 1.0 + 1.0 = 5.0 val.)

Consider the audio codec specified in MPEG-1 Audio Layer 3.

a) What is a reasonable coding rate to code a mono audio signal if each audio channel is sampled at 44 kHz and the codec above is used ? (R: 58.67 kbit/s)

b) Explain without computations how would the coding rate above change if the sampling frequency is halved and the signal becomes stereo ? (R: Stay the same)

c) Indicate two coding effects that may happen if the audio encoder includes a psychoacoustic model which is ´conservative' in the sense of using lower than appropriate hearing thresholds associated to the audio frequency masking effects ?

d) Indicate two coding effects that may happen if the audio encoder includes a psychoacoustic model which is 'optimistic' in the sense of using higher than appropriate hearing thresholds associated to the audio frequency masking effects ?

e) Considering the varied composition of a symphonic orchestra, how would change the quality of the audio subjective experience offered to the user if the audio bandwidth is successively enlarged starting from half the full bandwidth (22 kHz) while always adopting a sampling rate which respects the sampling theorem ? (R: Quality of user experience would successively increase)

IV (1.0 + 1.0 + 1.0 + 1.0 + 1.0 = 5.0 val.)

Consider the H.264/AVC video coding standard.

a) What is the key trade-off when selecting between the Intra16×16 and Intra4×4 coding modes ?

b) What is the key trade-off when selecting a macroblock partition for motion compensation with more or less motion vectors ?

c) What is the key trade-off when selecting the distance between the Intra frames in the temporal prediction structure, i.e. the N value ?

d) What is the key trade-off when selecting the number of (past and future) reference frames to have available in the buffer to be used for predictions ?

e) What is the key trade-off when selecting between CABAC and CAVLC for entropy coding ?