MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2018/2019 – 2nd Semester, Responsible: Prof. Fernando Pereira

1st Exam – 17th June 2019 (Monday), 8am

The marks should be out before **17th June (Monday), 6pm** at the CMul Web page and the exam checking session will on the **18th June (Tuesday), 10am** in room 0.17.

The exam is **3 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. ***Good luck !***

I (0.5 + 0.5 + 0.5 + 0.5 + 0.5 val. = 2.5 val.)

a) How many signal components are needed to see a gray image in a black and white television ? Which ones ? (R: one, luminance)

b) How many signal components are needed to see a gray image in a colour television ? Which ones ? (R: 3, RGB or luminance and chrominances)

c) How many signal components are needed to see a colour image in a colour television ? Which ones ? (R: 3, RGB or luminance and chrominances)

d) In total, are there more samples or pixels in a black and white image ? Why ? (R: the same)

e) Are there more luminance samples or total chrominance samples in a 4:2:2 colour image ? Why ? (R: the same)

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

Consider the JPEG standard to code 4:2:0 photographic images.

a) How many times would the number of luminance blocks increase if the color subsampling is changed to 4:2:2 ? And the total number of chrominance blocks for the same change ? (R: Same; doubles)

b) Determine the average number of bits per luminance sample that has to be spent to code this type of image if a codec with a luminance compression factor of 25 and a chrominances compression factor of 15 is used. (R: 0.32 bit/sample)

c) How many bits would cost a color image with 576×720 pixels in PCM ? (R: 4 976 640 bits)

d) How many blocks do you have to code for a color image with 576×720 pixels ? (R: 9720 blocks)

e) How many DCT coefficients would you have to code for the luminance component of an image with 576×720 pixels if the image has a single uniform color ? (R: 6480)

f) Determine the maximum number of DCT coefficients that may be coded for each luminance block of an image with 576×720 pixels if each coefficient costs, on average, 4 bits for the luminance and 5 bits for the chrominance and a maximum total number of bits of 400000 is desired; consider that luminance blocks always use 1 coefficient more that each chrominance block and additionally consider that the EOB (End of Block) word costs 3 bits. (R: 9)

III (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 = 3.0 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 maximum number of bits that the second frame may produce is 15000 and the buffer size is 12000, how many bits did the first frame produce ? (R: 22600 bit)

b) Assuming that the first, second and third frames produce each 15000 bits each, what is the minimum size of the buffer? (R: 6600 bit)

c) Assuming that the first frame produces 18000 bits and the buffer size is 10000 bits, what is the maximum number of bits that the third frame may produce ? (R: 22800 bit)

d) Assuming that the buffer size is 14000, what is the maximum number of bits that the first frame may produce if the encoder is infinitely fast, thus producing all the bits for each frame instantaneously ? (R: 14000 bit)

e) Assuming that the buffer size is 10000, what is the maximum number of bits that the second and third frames may produce ? (R: 22800 bit)

f) Assuming that the buffer is full when the 5^{th} frame starts to be encoded, what is the maximum number of bits that it may produce? (R: 12800 bit)

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

Consider the MPEG-1 Audio standard.

a) Determine the coding rate for mono audio content with a 20 kHz bandwidth and the usual number of bit/sample if coded with a Layer 3 codec to reach CD transparent quality. How much would the rate vary in percentage if the signal becomes stereo ? (R: 53.333 kbit/s; 100%)

b) What could happen if in a Layer 3 codec only the long MDCT window size is used ? Why ?

c) What could happen if in a Layer 3 codec no MDCT window overlapping is used ? Why ?

d) Considering the varied composition of a jazz orchestra, how would change the subjective quality assessment associated to the music experience if the full audio bandwidth is always used while successively increasing the initial sampling rate from a value which starts being 1.5 times the full bandwidth and ends being 3 times the full bandwidth? Why ? (R: Quality increases)

e) What would happen in terms of subjective quality assessment if the audio bandwidth is now increased from half the full bandwidth to full bandwidth with the sampling rate increasing in the same proportion starting from a value which is the double of the bandwidth ? Why ? (R: Quality increases)

V (1.0 + 1.0 + 1.0 + 0.5 + 0.5 = 4.0 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 4 minutes clips in a disk. The maximum access speed to the disk is 80 Mbit/s. The clips have 4K resolution with the following characteristics: 3840 x 2160 (Y), 4:2:2, 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 40 and 45 for the luminance and chrominances, respectively, determine the maximum access time for an image knowing that the compression factors for critical frames are 20% lower than average. (R: 61.12 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: 346.26 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)

d) Determine, justifying, which coding solution would you propose to your client if a maximum random access requirement of 100 ms is put forward together with the requirement of maximizing the number of clips stored in the disk. (R: JPEG)

e) How many full video clips would you be able to store in the disk for the JPEG solution if the disk has a capacity of 10 TByte (10^{12}). (R: 3404 full clips)

VI (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 = 2.5 val.)

Consider the video codec specified in H.264/AVC standard.

a) What may be service implications of the fact that H.264/AVC spends “about 50% less rate for the same perceptual quality regarding previous existing standards” (2 implications) ?

b) Why is it appropriate to say that H.264/AVC “does NOT allow to guarantee any minimum level of quality” ? (R: mainly because encoders are not normative)

c) How does H.264/AVC try to overcome the limitations of using a translational motion model ? (R: variable block size for motion compensation)

d) What is the main goal of using well selected half- and quarter-sample interpolation filters ? (R: make better temporal predictions)

e) What are the main positive and negative impacts of using multiple reference frames ? (R: better prediction and better RD performance versus more memory and more computational complexity)