COMUNICAÇÃO MULTIMÉDIA

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2014/2015 – 2nd Semester, Responsible: Prof. Fernando Pereira

2nd Exam – 3rd July 2015 (Friday), 9am

The marks should be out before **4th July (Saturday), 8pm** at the CMul Web page and the exam checking session will on the **6th July (Monday), 5pm** in room 0.32.

The exam is **<u>3 hours long</u>**. 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 + 1.0 + 0.5 + 1.0 = 3.0 val.)

Consider a broadcasting service …

a) What is the main performance target of an objective image/video quality metric ?

b) Is it feasible to use the PSNR to assess the video quality at the consumer premises/house ? Why ? (R: No as originals are not available)

c) Why doesn't the video PSNR necessarily guarantee a good assessment of the video quality? (R: Does not consider the Human Visual System properties)

d) What does typically consume more bitrate, data or metadata ? Why ? (R: Data)

II (1.0 + 1.0 + 1.0 + 0.5 val. = 3.5 val.)

Make some quick computations while neglecting the audio and the channel coding bitrates …

a) Assume that you have available a bandwidth slot of 8 MHz. How many standard definition MPEG-2 Video coded TV channels may you typically insert in this bandwidth with acceptable quality if a 64-QAM modulation is used ? (R: 12)

b) What would be the answer to the same question in a) if the H.264/AVC and the HEVC standards were used to code the video ? (R: 24 and 48)

c) What would be the answer to the same question in a) if each TV channel is now a MVC coded stereo channel where the dependent view is coded with the minimum rate necessary to guarantee the same stereo perceptual quality of a (rate equivalent) H.264/AVC simulcasting coding solution where both stereo channels use the same bitrate ? (R: 19)

d) What would be the answer to the same question in c) if a rate equivalent HEVC simulcasting coding solution (and not a MVC solution) is used ? (R: 24)

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 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 12800 bits, what is the maximum number of bits that the first frame may produce ? (R: 19200 bits)

b) Assuming that the first frame produces 20000 bits, what is the minimum size of the buffer ? (R: 13600 bits)

c) Assuming that the first frame produces 10000 bits and the buffer size is 12800 bits, what is the maximum number of bits that the second frame may produce ? (R:15600 bits)

d) Assuming that the first frame produces 15000 bits and the buffer size is 12800 bits, what is the maximum number of bits that the third frame may produce ? (R: 17000 bits)

e) Assuming that the buffer size is 12800, what is the advisable initial visualization delay ? (R: 300 ms)

f) Assuming that the buffer size is 12800, what is latest time the full set of bits for the first frame may be received at the decoder ? (R: 300 ms)

IV (1.0 + 1.0 + 1.0 + 0.5 + 1.0 = 4.5 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 with 10 TByte (10^{12}) of capacity. 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 Video)

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. (R: 3404 clips)

V (0.5 + 1.0 + 0.5 + 1.0 + 0.5 = 3.5 val.)

Consider an MPEG-1 Audio coding solution.

a) What is the main audio characteristic justifying to use a bit depth higher than for speech ? (R: Dynamic range)

b) In which two ways does the frequency masking effect contribute to reduce the audio coding rate ?

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

d) Would you prefer listen to 44 kHz sampling frequency and 16 bit/sample **OR** 96 kHz sampling frequency and 12 bit/sample audio ? Why ? (R: First solution)

e) What is the main reason to increase the size of the audio frames in MPEG-1 Audio Layer 2 regarding Layer 1?

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

Consider a DVB-x2 system.

a) Indicate a major difference between the coding of a B-macroblock in H.264/AVC versus MPEG-2 Video.

b) Why does H.264/AVC include a deblocking filter in the prediction loop and not outside the loop as for MPEG-2 Video ?

c) How can you control the channel coding module in terms of associated delay ? (R: Size of BCH+LDPC block)

d) In DVB-S2, what is the major advantage of APSK regarding an alternative QAM solution with the same number of symbols ?

e) Why does DVB-C2 adopt OFDM considering that cable transmissions do no suffer from multi-path reception problems ?