MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2017/2018 – 2nd Semester, Responsible: Prof. Fernando Pereira

1st Exam – 18th June 2018 (Monday), 8am

The marks should be out before **19th June (Tuesday), 12am** at the CMul Web page and the exam checking session will on the **19st June (Tuesday), 2pm** in room 1.4.

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 + 1.0 + 1.0 val. = 3.5 val.)

a) How many signal components are needed to see a gray image in a black and white television ? Which ones ? (R: 1, luminance)

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

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

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

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

Consider the JPEG standard to code photographic images with a 576×720 luminance resolution, 4:2:0 color subsampling and 8 bit/sample.

a) How many total pixels, samples and blocks exist in this type of image. (R: 414720 pixels, 622080 samples, 9720 blocks)

b) Determine the average price in bits (considering both the luminance and the chrominances) to code 100 pixels in this type of image if a codec with a luminance compression factor of 15 and a chrominances compression factor of 20 is used. (R: 73.3 bit)

c) Determine the total number of bits that have to be spent to code the luminance component of an 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 3 bits. (R: 149040 bit)

d) What does it mean saying that the entropy coder exploits the signal statistical redundancy ?

e) What JPEG coding tool typically brings undesirable numerical representation problems ? Why ?

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 bit)

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

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 bit)

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 bit)

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 (0.5 + 0.5 + 0.5 + 2.0 + 0.5 = 4.0 val.)

Suppose that you are contacted by an advertising company to design a multimedia digital storage system. The maximum access speed to the disk is 30 Mbit/s. The clips have HDTV resolution, this means 1920×1152 (Y) and 960×1152 (Cr, Cb) at 25 Hz. Assume that you have at your disposal providing the required video quality:

1. a JPEG coding solution with a compression factor of 25 for both the luminance and chrominances
2. a MPEG-2 Video coding solution with the following 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

a) If your client asks for the most compression efficient frame level random access solution (meaning that each frame should be independently accessed), what solution would you offer him/her from those above ? (R: MPEG-2 Video with only I frames as it has better compression factors than JPEG)

b) If your client asks for the most compression efficient, very low complexity coding solution, what solution would you offer him/her from those above ? (R: MPEG-2 Video with only I frames as it has better compression factors than JPEG)

c) If your client asks for a coding solution with best interoperability with the current image-coding ecosystem, what solution would you offer him/her from those above ? (R: JPEG)

d) If your client asks for the most compression efficient coding solution with a maximum access time per image below 400 ms, what solution would you offer him/her from those above if M=2 had to be used for the MPEG-2 Video solutions (if relevant, specify the GOP size) ? (R: MPEG-2 Video with N=24)

e) If you could have available a third coding solution, notably to increase the compression efficiency, what solution would you like to have at your disposal ? Why ? (R: H.264/AVC has it offers better compression efficiency, notably around 50% the rate for the same perceptual quality)

V (0.5 + 0.5 + 1.0 + 0.5 + 0.5 = 3.0 val.)

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

a) What is a reasonable rate required to code a stereo signal with 44kHz sampling rate ? (R: 117.333 kbit/s)

b) What does it happen if the ‘spatial integrity’ is lost when coding a stereo audio signal ?

c) Indicate 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 ?

d) 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?

e) 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 d) ?

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

Consider 3D video systems and concepts.

a) In what consists the so-called accommodation-vergence conflict ?

b) What and why is the typical baseline for a stereo camera ? (R: 65 mm)

c) What is the basic principle involved in frame compatible stereo formats ?

d) What type of redundancy may a multiview video encoder exploit that may not be exploited by a regular (mono) video encoder ? (R: inter-view redundancy)

e) What does it mean creating a disparity-compensated prediction ?

f) What is the basic idea behind depth-image-based rendering ?