MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2016/2017 – 2nd Semester, Responsible: Prof. Fernando Pereira

1st Exam – 19th June 2017 (Monday), 8am

The marks should be out before **20th June (Tuesday), 8pm** at the CMul Web page and the exam checking session will on the **21st June (Wednesday), 11.30am** in room 0.16.

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 + 0.5 val. = 3.0 val.)

Consider digital multimedia content.

a) Which are the 3 main parameters determining the uncompressed rate of a digital video signal ? (R: frame rate, spatial resolution and nº bit/sample)

b) Why is a YUV video representation more friendly for compression than a RGB video representation ?

c) Indicate 3 characteristics of the human vision that are critical to limit/define the uncompressed rate required to represent digital video.

d) What does the human visual Contrast Sensitivity Function express ? What may be its impact when coding an image ?

e) Considering the varied instrumental composition of a symphonic orchestra, how would change the subjective impression created by the sound if the bandwidth is successively reduced/filtered while not reducing the sampling rate ? (R: reduced quality associated to the elimination of higher frequencies for the instruments, even if the sampling rate is unnecessarily high for the bandwidth)

II (1.0 + 0.5 + 0.5 + 0.5 + 0.5 val. = 3.0 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, 622080, 9720)

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 numerical representation problems ? Why ? (R: DCT)

III (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 1.0 = 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 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)

g) Indicate one type of motion for which this codec may be especially efficient and another type of motion for which this codec is less efficient. Explain why. (R: translations and rotations)

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

Consider a ITU-T H.261 videophone communication working with a video signal at QCIF spatial resolution (176×144 samples for the luminance, 4:2:0 chrominance subsampling, with 8 bit/sample), 10 Hz. Assume that the average global (luminance and chrominances) compression factor (measured over all the macroblocks in the image), without any external constraints in terms of coding modes is 25 (with header bits included).

a) Explain what happens (and why) if, when the two videophones try to establish the communication, they find that one of them has CIF spatial resolution capabilities and the other only has QCIF resolution capabilities.

b) Assuming that for each frame, on average, only 50 macroblocks generate code bits (the remaining ones are so similar to the previous image that no update is needed), determine the average overall compression factor (including luminance and chrominances) measured over the macroblocks which effectively generate code bits. (R: 12.6)

c) For the situation in b), assume that to guarantee a higher error protection, one out of each 25 macroblocks spending bits is necessarily coded in intra mode. Assuming that the intra coding mode has a compression factor at the macroblock level which is half of the compression factor determined in b), determine what would be the global compression factor (considering all the macroblocks in the image) corresponding to this situation. (R: 24.04)

d) If a transmission rate of 40 kbit/s is used, what would be maximum number of bits that the **second frame** may spend if the first frame spends 6500 bits and a maximum acquisition-visualization delay of 200 ms would be required ? Assume that the encoder generates the bits for each frame uniformly in the time period between the acquisition moments of each two successive frames. (R: 5500 bit)

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

Consider the MPEG-1 and MPEG-2 Audio standards.

a) Determine the coding rate for stereo audio content with a 22 kHz bandwidth and the usual number of bit/sample if coded with a Layer 3 codec to reach CD transparent quality. How would the rate vary in percentage if the sampling rate becomes 48 kHz and mono audio is used. (R: 117333.3 bit/s; -45%)

b) What does it mean saying that audio does not have a universal source production model ? Why is this different for speech ?

c) Why does the Layer 3 codec use the MDCT with an overlapping window ? How is this overlapping applied ?

d) Why does the Layer 3 codec use the MDCT with a varying size window ? How is this varying size window applied ?

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

Consider a DVB based digital TV system.

a) Knowing that a DVB solution may initially 'insert' 10 Mbit/s of source rate in a 8 MHz bandwidth channel, determine what is the final *channel coding ratio* knowing that the total transmitted rate is always 3 times higher than the source rate and the modulation is made 4 times more efficient in the final solution. (R: 1/3)

b) Knowing that a DVB solution may 'insert' 10 Mbit/s of source rate in a 8 MHz bandwidth channel, what is the *total transmitted bitrate* if all the system parameters stay the same with the exception of the channel coding ratio that goes from 1/3 to 1/2 and the modulation that goes from 16-PSK to 16-QAM. (R: 30 Mbit/s)

c) What positive and negative happens in a DVB-T context if the modulation efficiency of each carrier is reduced while keeping all the remaining parameters, notably the total number of carriers ?

d) What positive and negative happens in a DVB-T context if the guard interval is made longer ?

e) What trade-off led to the definition of the COFDM 2k and 8k modes in DVB-T ?