MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2015/2016 – 2nd Semester, Responsible: Prof. Fernando Pereira

2nd Exam – 28th June 2016 (Tuesday), 9am

The marks should be out before **29th June (Wednesday), 8pm** at the CMul Web page and the exam checking session will on the **30th June (Thursday), 5.30pm** in room 0.25.

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 (1.0 + 0.5 + 1.0 = 2.5 val.)

Consider the coding of digital images.

a) Assuming the most typical number of bits per sample, what would be the compressed rate cost of one luminance sample and of one chrominance sample if the compression factors for the luminance and chrominance are 20 and 40, respectively ? (R: 0.4 and 0.2 bit/sample)

b) How many bits would cost, on average, a JPEG luminance block for the situation described in a) ? (R: 25.6 bit/block)

c) Assuming that 280000 bits are available to code a 4:4:4 image for the situation in a), what would be its maximum horizontal resolution knowing that its vertical resolution is 500 lines ? (R: 700 sample/line)

II (0.5 + 1.0 + 1.0 + 0.5 + 1.0) = 4.0 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), 10 Hz. Assume that the average global (luminance and chrominances) compression factor (measured over all the macroblocks in the image), without any *forced 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. (R: Communication in QCIF)

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 (luminance and chrominances) compression factor measured over the macroblocks which effectively generate code bits. (R: 12.62)

c) For the situation in b), assume that to guarantee a higher error resilience, 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 that is half of the compression factor determined in b), determine for this new situation what would be the *global* (luminance and chrominances) compression factor measured over all the macroblocks in the image. (R: 24.04)

d) If a transmission rate of 40 kbit/s is used, what would be maximum number of bits that a frame may spend if 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 pair of two successive frames. (R: 8000 bit)

e) For the situation in d), determine the maximum number of bits that the second frame could spend knowing that the first frame has spent 6500 bits. (R: 5500 bit)

III (1.0 + 1.0 + 1.0 + 1.0 = 4.0 val.)

Consider the audio coding solutions as adopted in the MPEG-1 Audio standard.

a) Explain what do the so-called *threshold of hearing* and *threshold of pain* express in terms of human auditory perception. What is the impact of these thresholds in terms of audio coding ?

b) Explain what is the so-called *masking effect* in audio perception. What is the difference between frequency and temporal masking ?

c) What is the basic idea behind *perceptual audio coding* ? Why does audio coding exploit the characteristics of the human auditory system and not also of the human vocal system like in speech coding ?

d) Identify and explain the 2 main ways to perform *frequency driven audio coding* as adopted in the MPEG-1 Audio standard. Which main audio coding solution does simultaneously exploit these two ways ?

IV (1.0 + 1.5 + 1.0 = 3.5 val.)

Consider that your company is contacted to design a videoconference system between the various main locations of a bank. The spatial resolution is CIF (352×288 luminance samples), 4:2:0, at 12.5 Hz, with the usual number of bits per sample. Assume that you have available, offering the target video quality, two solutions:

1. **H.261 based solution** with average compression factors of 25 and 35 for the luminance and chrominance, respectively; the critical compression factors (for the images spending more bits) are 20 and 25 for the luminance and chrominances, respectively.
2. **MPEG-2 Video based solution** with N = M = 3 with average compression factors of 25 and 35 for the luminance and chrominance, respectively, for the I frames, and 30 and 45 for the luminance and chrominances, respectively, for the P and B frames. The critical compression factors are 75% of the average compression factors.

Assume that the transmission rate is always the same as the coding rate.

a) Determine the bitrate and acquisition-visualization delay for the H.261 based solution. (R: 550326.9 bit/s and 103 ms)

b) Determine the bitrate and acquisition-visualization delay for the MPEG-2 Video based solution. (R: 483815.6 bit/s and 400 ms)

c) Assuming that your client always pretends to minimize the transmission rate, what solution from above would you select as a function of the acquisition-visualization delay requirement defined by the client?

V (0.5 + 1.0 + 0.8 + 0.7 = 3.0 val.)

Consider a DVB based digital TV system.

a) Knowing that a DVB solution may 'insert' 10 Mbit/s of source rate in a 8 MHz bandwidth channel, determine what is the *channel coding ratio* knowing that the transmitted rate is 3 times higher than the source rate. (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 *transmitted bitrate* if all the system parameters stay the same with the exception of the channel coding ratio that goes from 1/3 to 1/4 and the modulation that goes from 8-PSK to 16-QAM. (R: 40 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 trade-off led to the definition of the COFDM 2k and 8k modes in DVB-T ? What is the benefit of adding more COFDM modes as it has been done in DVB-T2 ?

VI (1.0 + 1.0 + 1.0 = 3.0 val.)

Consider the H.264/AVC video coding standard.

a) Considering interlaced content, for what type of content is *Frame Coding Type 1* more appropriate ? And *Frame Coding Type 2* ?

b) For what type of content are the *4×4 Intra prediction modes* more appropriate ? And the *16×16 Intra prediction modes* ?

c) For what type of content are the *bigger motion partition types* (e.g. 16×16) more appropriate ? And the *smaller motion partition types* (e.g. 4×4) ?