MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

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

1st Exam – 13th June 2016 (Monday), 9am

The marks should be out before **14th June (Tuesday), 2pm** at the CMul Web page and the exam checking session will on the **14th June (Tuesday), 5pm** in room 0.15.

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 + 1.0 + 0.5 + 1.0 + 1.0 val. = 4.5 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) Why is it reasonable to ask that an ideal transform for image compression compacts the information/signal energy as much as possible ?

b) Why is it reasonable to ask that an ideal transform for image compression creates coefficients which are as much as possible uncorrelated ?

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

d) 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)

e) Determine the maximum number of DCT coefficients that may be coded for each luminance block of an image if each coefficient costs, on average, 5 bits for the luminance and 4 bits for the chrominance and a maximum total number of bits of 400000 is desired; consider that luminance blocks always code 1 coefficient more that each chrominance block and additionally consider that the EOB (End of Block) word costs 3 bits. (R: 8 coefficients)

f) Indicate the main JPEG baseline encoder architectural element responsible for this type of coding being lossy. Explain why. (R: Quantizer)

II (0.5 + 0.5 + 0.5 + 0.5 + 2.0 val. = 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 rate of 64 kbit/s.

The video content to code is horizontally divided into **three equal parts;** however, while the top and bottom parts are fixed, the middle part is moving. Since the encoder processes sequentially the macroblocks, it is observed that all bits are uniformly generated in the central, one third part of the time interval that the encoder usually dedicates to encode each image. At the encoder, the bits wait for transmission in an output buffer.

Knowing that the first image has used 12800 bits, the second image 16000 bit, and the third image 6400 bits, determine:

a) The time instants at which the sender acquires the first, second and fourth images. (R: 0, 100 and 300 ms)

b) The time instants at which the sender finishes to produce the bits for the first, second and fourth images. (R: 100, 166 and 366 ms)

c) The time instants at which the sender finishes to send the bits for the first, second and third images. (R: 200, 450, 550 ms)

d) The time instants at which the receiver finishes to receive the bits for the first, second and third images. (R: 200, 450, 550 ms)

e) The time instants at which the display should show the first, third and fifth images assuming that the minimum acceptable acquisition-visualization delay is applied and the minimum acceptable buffer size is used. (R: 383, 583 and 783 ms)

III (0.5 + 2.0 + 0.5 = 3.0 val.)

Suppose that you are contacted by an advertising company to design a multimedia digital storage system. The company requires editing flexibility with a maximum access time per image below 400 ms and needs to store the largest number of 3.8 minutes clips in a disk with 500 TBytes of capacity. 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. Assuming that you have at your disposal providing the required video quality:

1. a JPEG coding solution with a compression factor of 30 for both the luminance and chrominances
2. a MPEG-2 Video coding solution with the following compression factors when M=2 is used:

· I frames: 20 and 25 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) Considering its native features, which of these two coding solutions is more suitable for editing purposes? Why ?

b) Determine, justifying, which coding solution should be proposed to your client ? Identify a major limitation of this proposed solution. (R: MPEG-2 Video)

c) How many full video clips would you be able to store in the disk with the two coding solutions above. (R: 594888 for JPEG and 924918 for MPEG-2 Video)

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

Consider the MPEG-1 Audio standard.

a) How many bits would you typically need to code a 5 minutes music clip, mono, with 20 kHz bandwidth and the typical number of bits /sample if using each of the 3 MPEG-1 Audio coding layers as the coding solutions to reach CD transparent audio quality ? (R: 48×10^6, 24×10^6 and 16×10^6)

b) Between MPEG1- Audio layers 1 and 2 which one has the largest coding delay ? Why ? And the highest encoding complexity ? Why ?

c) What would happen if you try decoding a MPEG-1 Audio layer 2 stream with a MPEG-1 Audio layer 3 decoder ? Why ?

d) In general, between the *block effect* and the *pre-echo effect* which one do you think may have the highest negative effect in terms of subjective quality if systematically poorly resolved ? Why ?

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

Consider the DVB standards.

a) Why are QPSK and 8-PSK typically used for television in DVB-S2 knowing that this standard accepts other modulations for satellite transmission ?

b) Considering that DVC-2 should not have to accommodate multipath effects, why does DVC-C2 recommend COFDM as the modulation to use ?

c) What is the advantage for DVB-T2 of including a larger set of OFDM modes in terms of number of carriers ?

d) Why is it normal to have a simulcasting period when introducing digital TV in a country which already has analogue TV ?

e) Knowing that a DVB solution may initially 'insert' 20 Mbit/s of source rate in a 8 MHz bandwidth channel using a channel coding ratio of 1/2, what is the transmitted source rate if all the system parameters stay the same with the exception of the modulation that goes from 16-QAM to 8-PSK. (R: 15 Mbit/s)

VI (0.7 + 0.7 + 0.6 = 2.0 val.)

Consider the H.264/AVC video coding standard.

a) By how much would the number of TV channels in a 8 MHz slot increase when using the H.264/AVC standard regarding the best previous standard coding solutions if all the technical elements stay the same ? Why ?

b) What is the main advantage of using H.264/AVC frame coding Type 2 regarding H.264/AVC frame coding Type 1 ?

c) What is the main consequence of using Intra prediction modes in terms of the Intra coding block energy to transmit?