MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2018/2019 – 2nd Semester, Responsible: Prof. Fernando Pereira

2nd Exam – 9th July 2019 (Tuesday), 3pm

The marks should be out before **12th July (Friday), 2pm** at the CMul Web page and the exam checking session will on the **12th July (Friday), 5pm** in room 0.23.

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 + 0.5 + 0.5 + 0.5 val. = 3.0 val.)

a) What is the difference between the bandwidth and the sampling frequency of a video signal ?

b) Indicate two benefits for Humans to have binocular vision.

c) What property relates the R, G and B values of a grey pixel ? Why ?

d) What are the chrominance (U and V) values of a grey pixel ? Why ?

e) Why do video services typically transmit chrominances (UV) with luminance and not directly primary colors (RGB) ?

f) Why do video services typically use 4:2:0 chroma subsampling ?

II (1.0 + 0.5 + 0.5 + 1.0 +0.5 = 3.5 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 (separately) 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) How many bits would cost, on average, a JPEG luminance image with resolution 576×720 for the situation described in a) ? (R: 165888 bit)

d) 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)

e) Assuming that the same 280000 bits are available to code a 4:2:0 image, and the same situation as in a) applies, how would the total number of pixels in the image change (same, increase or decrease) in comparison with the situation in d) ? (R: increase)

III (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 1.0 val. = 4.0 val.)

Consider a videotelephony communication using Recommendation ITU-T H.261. The video sequence is coded with a CIF spatial resolution (352×288 samples for the luminance, 4:2:0), a frame rate of 10 Hz and a constant bitrate channel of 64 kbit/s. The output buffer has a size of 12800 bits. The bits for each coded image are uniformly generated in the time between the acquisition of two images. Answer the following questions, always explaining.

a) What is the average number of bits that each image captured by the camera may spend ? (R: 6400 bit)

b) What is the maximum number of bits that each image captured by the camera may spend ? (R: 19200 bit)

c) What is the minimum number of bits that each image captured by the camera may spend ? (R: 0 bit)

d) What is the maximum number of bits that the third image may spend ? (R: 19200 bit)

e) What is the maximum number of bits that a image may spend if the buffer is full when the image starts to be coded ? (R: 6400 bit)

f) What is the maximum number of motion vectors that may be used to code a single image ? (R: 396)

g) Considering that a constant bitrate channel is used, what architectural element takes the most important decisions to allow the encoder controlling the number of bits spent per frame ? Why ? (R: Quantizer)

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 only with I frames)

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 only with I frames)

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)

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: 117.333 kbit/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 + 0.5 + 0.5 + 0.5 = 2.5 val.)

Consider the H.264/AVC video coding standard.

a) What is the number of macroblocks in an image with 576×720 luminance resolution ? (R: 1620)

b) What is the maximum number of motion vectors that a B-macroblock may have if no luminance blocks smaller than 8×8 samples are used for motion estimation ? Explain. (R: 8)

c) What is the maximum number of 4×4 integer DCT transforms that may have to be computed when coding a 4:2:0 macroblock ? Explain. (R: 24)

d) How many more Intra coding modes may be used if 4×4 and not 16×16 luminance blocks are used for the Intra prediction ? Explain. (R: 5)

e) What is a good estimation of the coding rate used to code a video with H.264/AVC, knowing that the same video spent 5 Mbit/s with MPEG-2 Video coding, if the same perceptual quality is targeted ? Explain. (R: 2.5 Mbit/s)