TÉCNICO LISBOA

MULTIMEDIA COMMUNICATION

INSTITUTO SUPERIOR TÉCNICO

Academic Year 2020/2021 – 2nd Semester, Responsible: Prof. Fernando Pereira

1st Exam – 23rd June 2021 (Wednesday), 11.30am

The marks should be out before **24th June (Thursday), 2pm** at the CMul Web page. If you want to check the exam scoring, please contact the course responsible by email **before 25th June (Friday)** to schedule a conversation.

The exam is **2.5 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. ***Boa sorte !***

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

a) Explain why any analogue multimedia signal becomes 'richer' when its bandwidth increases. (R. more frequencies and faster variations)

b) Explain why any digital multimedia signal becomes more 'complex' when the corresponding analogue multimedia signal increases its bandwidth. (R: higher sampling rate, memory, etc.)

c) Explain what mainly determines the number of bits per sample when digitizing an analogue multimedia signal, e.g. speech, audio, video. Why is this number different for speech and audio ? (R: transparency regarding analogue version; different dynamic range)

d) For video, what is the key factor determining a proper frame rate ? What happens to the user experience if the frame rate is too low ? (R: illusion of motion; hiccup effect)

e) For video, how should the spatial resolution change if the display size increases OR the viewing distance reduces ? (R: must increase)

II (1.0 + 1.0 + 1.0 + 1.0 val. = 4.0 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 times would the number of luminance blocks increase if the color subsampling is changed to 4:2:2 ? And the total number of chrominance blocks for the same change ? (R: does not change; doubles)

b) Determine the average number of bits per pixel 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: 0.587 bit/px)

c) Determine the chrominances compression factor if it is required that the rate cost for the luminance of one pixel is the same as for the two chrominance components of one pixel if the luminance compression factor is 25, assuming the color subsampling is now 4:4:4. (R: 50)

d) Determine the maximum number of DCT coefficients that may be coded for each luminance block of an image if each coefficient costs, on average, 4 bits for the luminance and 3 bits for the chrominance and a maximum total spending of 350000 bits for one image is desired; consider that luminance blocks always code 2 coefficients more that each chrominance block and additionally consider that the EOB (End of Block) word costs 3 bits. (R: 9)

III (0.5 + 0.5 + 1.0 + 0.5 + 1.0 + 0.5 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 128 kbit/s. The output buffer has a size of 20000 bits. The bits for each coded image are uniformly generated in the time between the acquisition of two images.

a) What is the average number of bits that a single frame may spend ? (R: 12800 bit)
b) What is the maximum number of bits that a single frame may spend ? (R: 32800 bit)
c) What is the maximum number of bits that a frame may spend if the previous frame spent 15000 ? (R. 30600 bit)
d) What is the maximum number of bits that a single frame may spend if the buffer is half-full when the frame is acquired ? (R: 22800 bit)
e) What are the maximum and minimum numbers of motion vectors that may be used to code a single frame ? (R: 0; 396)
f) Considering that a constant bitrate channel is used, which architectural element mostly allows the encoder controlling the quality of the frame ? Why ? (R: Quantizer)

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

Assume that you are contacted by a company to design a digital storage system for video clips. The company requires some editing flexibility and needs to store the largest number of 4 minutes clips in a disk. The maximum access speed to the disk is 80 Mbit/s. The clips have 4K resolution with the following characteristics: 3840 × 2160 (Y), 4:2:2, 10 bit/sample at 25 Hz.

a) Assuming that you have at your disposal, providing the required video quality, a JPEG coding solution with average compression factors of 40 and 45 for the luminance and chrominances, respectively, determine the maximum access time for a frame knowing that the compression factors for critical frames are 20% lower than average. (R: 61.2 ms)
b) Assuming now that you have at your disposal, providing the required video quality, a MPEG-2 Video coding solution with N=12 and M=4 with the following average 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

   Determine the maximum access time for a frame knowing that the compression factors for critical frames are 25% lower than average. (R: 346.26 ms)
c) Determine, justifying, which coding solution would you propose to your client if the target is only to maximize the number of clips stored in the disk.
d) Determine, justifying, which coding solution would you propose to your client if a maximum random access requirement of 100 ms is put forward together with the requirement of maximizing the number of clips stored in the disk. (R: JPEG as < 100 ms)
e) How many full video clips would you be able to store in the disk for the JPEG solution if the disk has a capacity of 10 TByte (10^{12}). (R: 3404)

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

Consider the H.264/AVC video coding standard.

a) Compute numerically the typical overall compression factor reached for (color) SD TV (576×720 pixels, 4:2:0) at 25 Hz with H.264/AVC coding. (R: 62.2)

b) From the two H.264/AVC lossless coding modes available, which one typically reaches higher compression ? Why ? (R: Lossless Coding Mode)

c) From the two H.264/AVC frame coding modes available for interlaced content, which one is more adaptive? Why ? (R: Type 2)

d) What processes does it involve coding the H.264/AVC motion vectors in a predictive way ? What is the main goal of predictive coding from a performance point of view ? (R: prediction and residue computation)