COMUNICAÇÃO DE ÁUDIO E VÍDEO

INSTITUTO SUPERIOR TÉCNICO

Year 2018/2019 – 1st Semester, Responsible: Prof. Fernando Pereira

2nd Exam – 29th January 2019, 8am (Tuesday)

The marks should be out before **29th January (Tuesday), 8pm** at the CAV Web page and the exam checking session will be on the **30th January (Wednesday), 11am** in room LT4.

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 type of retina cells is more important for night vision ? Why ?

b) Why it is commonly said that humans have trichromatic vision ?

c) What is the main impact of the human Contrast Sensitivity Function (CSF) in terms of image coding ?

d) What is the impact of stereo vision in terms of human field of view ?

e) If "it is possible to get a good illusion of motion up from 16-18 image/s, depending on the image content", why do European TV systems use 25 images per second ?

f) What does the visual acuity tell about the minimum required image resolution if the target viewing distance to the screen is reduced ?

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

Consider the JPEG standard to code 4:2:0 photographic images.

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: Same; doubles)

b) 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)

c) How many bits would cost a color image with 576×720 pixels in PCM ? (R: 4976640)

d) How many blocks do you have to code for a color image with 576×720 pixels ? (R: 9720)

e) How many DCT coefficients would you have to code for the luminance component of an image with 576×720 pixels if the image has a single uniform color ? (R: 6480)

f) Determine the maximum number of DCT coefficients that may be coded for each luminance block of an image with 576×720 pixels if each coefficient costs, on average, 4 bits for the luminance and 5 bits for the chrominance and a maximum total number of bits of 400000 is desired; consider that luminance blocks always use 1 coefficient 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 + 0.5 + 0.5 + 0.5 + 0.5 = 3.0 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 128 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 maximum number of bits that the second frame may produce is 15000 and the buffer size is 12000, how many bits did the first frame produce ? (R: 22600)

b) Assuming that the first, second and third frames produce each 15000 bits each, what is the minimum size of the buffer ? (R: 6600)

c) Assuming that the first frame produces 18000 bits and the buffer size is 10000 bits, what is the maximum number of bits that the third frame may produce ? (R: 22800)

d) Assuming that the buffer size is 14000, what is the maximum number of bits that the first frame may produce if the encoder is infinitely fast, thus producing all the bits for each frame instantaneously ? (R: 14000)

e) Assuming that the buffer size is 10000, what is the maximum number of bits that the second and third frames may produce ? (R: 22800)

f) Assuming that the buffer is full when the 5th frame starts to be encoded, what is the maximum number of bits that it may produce? (R:12800)

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

Consider the MPEG-1 Audio standard.

a) Determine the coding rate for mono audio content with a 20 kHz bandwidth and the usual number of bit/sample if coded with a Layer 3 codec to reach CD transparent quality. How much would the rate vary in percentage if the signal becomes stereo ? (R: 53,333 kbit/s)

b) What could happen if in a Layer 3 codec only the long MDCT window size is used ? Why ?

c) What could happen if in a Layer 3 codec no MDCT window overlapping is used ? Why ?

d) Considering the varied composition of a jazz orchestra, how would change the subjective quality assessment associated to the music experience if the full audio bandwidth is always used while successively increasing the initial sampling rate from a value which starts being 1.5 times the full bandwidth and ends being 3 times the full bandwidth? Why ?

e) What would happen in terms of subjective quality assessment if the audio bandwidth is now increased from half the full bandwidth to full bandwidth with the sampling rate increasing in the same proportion starting from a value which is the double of the bandwidth ? Why ?

V (1.0 + 1.5 + 0.5 = 3.0 val.)

Consider that your company is contacted to design a videoconference system between the various European premises of a pharmaceutical company. Your client pretends to use transmission lines with the minimum possible bitrate, providing the target quality, with a acquisition-visualization delay that should not exceed 250 ms. The spatial resolution is CIF (352×288 luminance samples), 4:2:0, at 10 Hz, with the usual number of bits per sample. Assume that you have available, offering the target quality, two solutions:

1. *H.261 based solution with average compression factors of 20 and 30 for the luminance and chrominance, respectively; the critical compression factors (for the images spending more bits) are 10 and 15 for the luminance and chrominance, respectively.*
2. *MPEG-2 Video based solution with N = M = 2 with average compression factors of 20 and 30 for the luminance and chrominance, respectively, for the I frames, and 30 and 45 for the luminance and chrominance, respectively, for the P and B frames. The critical compression factors are 80% 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: 540.672 kbit/s; 200 ms)

b) Determine the bitrate and acquisition-visualization delay for the MPEG-2 based solution. (R: 450.56 kbit/s; 350 ms)

c) Which solution would you recommend to your client considering the requirements defined above ? (R: H.261)

VI (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 = 3.0 val.)

a) What is the retinal disparity ?

b) Why doesn't the accommodation-vergence conflict happen in real life ?

c) What would you do to the baseline of a stereo camera if you wanted to intensify the final depth effect ?

d) Explain why are occlusions a good hint/cue for 3D/depth perception? Why is it a monocular cue ?

e) What is the main advantage of exploiting interview redundancy as in conventional stereo coding ?

f) Why is it problematic to have large coding errors at the sharp edges corresponding to object borders in a depth map ?