COMUNICAÇÃO DE ÁUDIO E VÍDEO

INSTITUTO SUPERIOR TÉCNICO

Year 2019/2020 – 1st Semester, Responsible: Prof. Fernando Pereira

1st Exam – 10th January 2020, 8am (Friday)

The grading should be out before **10th January (Friday), 7pm** at the CAV Web page and the exam checking session will be on the **13th January (Monday), 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 + 1.0 val. = 3.5 val.)

a) Why do video services typically transmit luminance and chrominances and not RGB components ?

b) How is it possible to create gray images in a color television ?

c) What would you do in terms of display if receiving a streaming service where one of the chrominance components is missing (hopefully only for some limited time) ? (R: Just show luminance)

d) What type of decoded block does it result if the block is coded only with the DC DCT coefficient and another DCT coefficient on the diagonal of the transformed block ?

e) What type of decoded block does it result if the block is coded only with the DC DCT coefficient and only another DCT coefficient on first line OR first row of the transformed block ?

f) Why does it make sense to use a non-uniform quantizer, notably for speech and audio signals ? How do you relate this with the Weber's law ?

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

Consider the JPEG standard to code photographic images with a 576×720 luminance resolution, 4:4:4 color subsampling and 8 bit/sample.

a) How many pixels, samples and blocks exist in this type of image. (R: 414720, 1244160, 19440)

b) Determine the price in bits per pixel if a codec with a luminance compression factor of 20 and a chrominances compression factor of 10 is used. (R: 2 bpp)

c) Determine the price in bits for a full image if a codec with a luminance compression factor of 25 and a chrominances compression factor of 15 is used. (R: 575078,4 bit)

d) Determine the price in bits to code a gray scale version of a full image with only 128 levels of gray if the same compression factors as in c) are used. (R: 116121,6 bit)

e) Determine the total number of bits that have to be spent to code only the luminance component of one image if an average number of 5 DCT coefficients are coded per block and each coefficient costs, on average, 3 bits; additionally consider that the EOB (End of Block) word costs 4 bits. (R: 123120 bit)

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 first and second frames produce 20000 bits each, what is the minimum size of the buffer ? (R: 14400 bit)

b) Assuming that the buffer with size 19000 has 10000 bits at 100 ms, what is the maximum number of bits that the second frame may produce ? (R: 21800 bit)

c) Assuming that the first frame produces 22000 bits and the buffer size is 10000 bits, what is the maximum number of bits that the second frame may produce ? (R: 13600 bit)

d) Assuming that the first frame produces 22000 bits and the buffer size is 20000 bits, what is the maximum number of bits that the third frame may produce ? (R: 32800 bit)

e) Assuming that the buffer size is 25000, what is the advisable initial visualization delay ? (R: 295,31 ms)

f) Assuming that the buffer size is 12000, what is latest time the full set of bits for the second frame may be received at the decoder ? (R: 293,75 ms)

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

Consider the MPEG-1 Audio standard.

a) Determine the coding rate for mono audio content with a 40 kHz sampling rate and the usual number of bit/sample if coded with a Layer 2 codec to reach CD transparent quality. How would the rate vary **in percentage** if the sampling rate becomes 48 kHz and stereo audio is used. (R: 80 kbit/s, 140%)

b) What does it mean saying that speech has a universal source production model ? Why is this different for audio ?

c) Why does the Layer 3 codec use the MDCT with a varying size window ? How is this varying size window applied ?

d) Why does the Layer 3 codec use the MDCT with an overlapping window ? How is this overlapping applied ?

e) Consider an audio signal with a full bandwidth of 22 kHz. Would you prefer to listen to a variation of this signal with a lower bandwidth and an appropriate sampling rate OR to a variation with the same bandwidth and a lower than Nyquist theorem sampling rate ? (R: Lower bandwidth and an appropriate sampling rate)

V (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, 103 ms)

b) Determine the bitrate and acquisition-visualization delay for the MPEG-2 Video based solution. (R: 483815,6 bit/s, 400 ms)

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

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

Consider the video codec specified in H.264/AVC standard.

a) What may be service implications of the fact that H.264/AVC spends “about 50% less rate for the same perceptual quality regarding previous existing standards” (2 implications) ?

b) Why is it appropriate to say that H.264/AVC “does NOT allow to guarantee any minimum level of quality” ?

c) How does H.264/AVC try to overcome the limitations of using a translational motion model ?

d) What is the main goal of using well selected half- and quarter-sample interpolation filters ?

e) What are the main positive and negative impacts of using multiple reference frames ?

f) What is a main practical impact (different from the past) of including in List 0 and List 1 the same (already decoded) frames although in a different order ?