COMUNICAÇÃO DE ÁUDIO E VÍDEO

INSTITUTO SUPERIOR TÉCNICO

Year 2016/2017 – 1st Semester, Responsible: Prof. Fernando Pereira

1st Exam – 12th January 2017, 8am (Thursday)

The marks should be out before **17th January (Tuesday), 7pm** at the CAV Web page and the exam checking session will be on the **18th January (Wednesday), 9.30pm** 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 + 1.0 + 0.5 val. = 3.0 val.)

Consider digital multimedia content.

a) Which are the 3 main parameters determining the uncompressed rate of a digital video signal ? (R: spatial resolution, frame rate and number of bits per sample)

b) Why is a YUV representation more friendly for compression than a RGB representation ?

c) Indicate 3 characteristics of the human vision that are critical to determine the uncompressed rate required to represent digital video.

d) What does the human visual Contrast Sensitivity Function express ? What may be its impact when coding an image ?

e) Considering the varied composition of a symphonic orchestra, how would change the subjective impression created by the sound if the bandwidth is successively reduced/filtered while not reducing the sampling rate ?

II (1.0 + 0.5 + 0.5 + 0.5 + 0.5 val. = 3.0 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 total pixels, samples and blocks exist in this type of image. (R: 414720 pixels; 622080 samples; 9720 blocks)

b) Determine the average price in bits (considering both the luminance and the chrominances) to code 100 pixels in this type of image if a codec with a luminance compression factor of 15 and a chrominances compression factor of 20 is used. (R: 73,3 bit)

c) Determine the total number of bits that have to be spent to code the luminance component of an image if an average number of 4 DCT coefficients are coded per block and each coefficient costs, on average, 5 bits; additionally consider that the EOB (End of Block) word costs 3 bits. (R: 149040 bit)

d) What does it mean saying that the entropy coder exploits the signal statistical redundancy ?

e) What JPEG coding tool typically brings numerical representation problems ? Why ?

III (0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 1.0 = 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 (constant) channel bitrate of 64 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 buffer size is 12800 bits, what is the maximum number of bits that the first frame may produce ? (R: 19200 bit)

b) Assuming that the first frame produces 20000 bits, what is the minimum size of the buffer ? (R: 13600 bit)

c) Assuming that the first frame produces 10000 bits and the buffer size is 12800 bits, what is the maximum number of bits that the second frame may produce ? (R: 15600 bit)

d) Assuming that the first frame produces 15000 bits and the buffer size is 12800 bits, what is the maximum number of bits that the third frame may produce ? (R: 17000 bit)

e) Assuming that the buffer size is 12800, what is the advisable initial visualization delay ? (R: 300 ms)

f) Assuming that the buffer size is 12800, what is latest time the full set of bits for the first frame may be received at the decoder ? (R: 300 ms)

g) Indicate one type of motion for which this codec may be especially efficient and another type of motion for which this codec is less efficient. Explain why. (R: Translations and rotations)

IV (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: -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 ?

V (0.9 + 1.0 + 1.0 + 0.6 = 3.5 val.)

Consider the development of an application similar to YouTube.

a) Indicate three main technologies required for the *free access* **representation** of the multimedia content involved ? For each of these technologies, define its main objective. (R: video coding, audio coding and metadata creation)

b) Considering the rate budget, provide reasonable estimates of the rate percentages associated to at least two of the main technologies referred in a), assuming that streaming is performed with standard definition video and stereo audio ? Justify your answer.

c) Considering the video component, what codecs would you support on the 'uploading' side and on the streaming side ? Consider both the cases of *live broadcasting* of a public event and *a posteriori streaming* of personal content. Justify your choices.

d) For the two cases defined in c), what type of video coding prediction structure would you recommend ? Why?

VI (0.5 + 0.5 + 0.9 + 0.8 + 0.8 = 3.5 val.)

Consider a DVB based digital TV system.

a) Knowing that a DVB solution may initially 'insert' 10 Mbit/s of source rate in a 8 MHz bandwidth channel, determine what is the final *channel coding ratio* knowing that the total transmitted rate is always 3 times higher than the source rate and the modulation is made 4 times more efficient in the final solution. (R: 1/3)

b) Knowing that a DVB solution may 'insert' 10 Mbit/s of source rate in a 8 MHz bandwidth channel, what is the *total transmitted bitrate* if all the system parameters stay the same with the exception of the channel coding ratio that goes from 1/3 to 1/2 and the modulation that goes from 16-PSK to 16-QAM. (R: 30 Mbit/s)

c) What positive and negative happens in a DVB-T context if the modulation efficiency of each carrier is reduced while keeping all the remaining parameters, notably the total number of carriers ?

d) What positive and negative happens in a DVB-T context if the guard interval is made longer ?

e) What trade-off led to the definition of the COFDM 2k and 8k modes in DVB-T ?