COMUNICAÇÃO DE ÁUDIO E VÍDEO

INSTITUTO SUPERIOR TÉCNICO

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

2nd Exam – 30th January 2017, 8am (Monday)

The marks should be out before **31st January (Tuesday), 8pm** at the CAV Web page and the exam checking session will be on the **1st February (Wednesday), 10am** 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 (0.8 + 0.5 + 0.7 + 1.0 = 3.0 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 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) 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)

d) Explain one positive and one negative consequence of increasing the block size (and the associated DCT transform) in JPEG from 8×8 to 16×16 samples.

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

a) What is the average number of bits that a single image may spend ? (R: 6400)

b) What is the maximum number of bits that a single image may spend ? (R: 19200)

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

d) What is the maximum number of bits that a single image may spend if the buffer is full when the image is acquired ? (R: 6400)

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

f) Considering that a constant bitrate channel is used, what architectural element mostly allows the encoder controlling the number of bits spent per frame ? Why ? (R: Quantizer)

III (0.5 + 0.5 + 2.0 + 0.5 = 3.5 val.)

Suppose that you are contacted by an advertising company to design a multimedia digital storage system. The company requires editing flexibility with a maximum access time per image below 400 ms and needs to store the largest number of 3.8 minutes clips in a disk with 500 TBytes of capacity. 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. Assuming that you have at your disposal providing the required video quality:

1. a JPEG coding solution with a compression factor of 30 for both the luminance and chrominances
2. a MPEG-2 Video coding solution with the following compression factors when M=2 is used:

- I frames: 20 and 25 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) Considering its native features, which of these two coding solutions is more suitable for editing purposes? Why ? (R: JPEG)

b) Considering its native features, which of these two coding solutions is more suitable for compression purposes? Why ? (R: MPEG-2 Video)

c) Considering the requirements defined above, determine, justifying, which coding solution should be proposed to your client ? (R: MPEG-2 Video)

d) If you could consider a third coding solution, which solution would you like to have at your disposal ? Why ?

IV (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: 550.37 kbit/s and 103 ms)

b) Determine the bitrate and acquisition-visualization delay for the MPEG-2 Video based solution. (R: 483.816 kbit/s and 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?

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

Consider the audio codec specified in MPEG-1 Audio Layer 3.

a) What is a reasonable rate required to code a stereo signal with 44kHz sampling rate ? (R: 117.333 kbit/s)

b) What does it happen if the 'spatial integrity' is lost when coding a stereo audio signal ?

c) Indicate two effects that may happen if the audio encoder includes a psychoacoustic model which is conservative in the sense of lowering the hearing thresholds associated to the audio masking effects ?

d) 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 initial sampling rate which respects the sampling theorem ?

e) What would happen differently from d) in terms of subjective impression if the initial sampling rate is also reduced in the same proportion of the bandwidth ?

VI (0.5 + 1.0 + 1.0 + 0.5 = 3.0 val.)

Consider a digital TV system for terrestrial broadcasting.

a) What is the main reason motivating the selection of a multi-carrier modulation for this type of system ?

b) What is the main difference between a FDM and a OFDM solution ? What is the main benefit of OFDM regarding FDM ?

c) What is the main characteristic of orthogonal carriers ? What is their main benefit ?

d) Is it good or bad or both using a longer guard interval ? Why ?