Information theory and coding – Image, video and audio compression –

exercises with some example solutions for supervisors

Markus Kuhn

Michaelmas 2004 – Part II

`http://www.cl.cam.ac.uk/Teaching/2004/InfoTheory/mgk/`

Exercise 1 Compare the quantization techniques used in the digital telephone network and in audio compact disks. Which factors to you think led to the choice of different techniques and parameters here?

Example solution:

Telephone system uses 8-bit logarithmic (A/μ-law) at 8000 Hz sampling frequency, audio CD uses 16-bit linear at 44.1 kHz sampling frequency. Input signal level can vary significantly with telephone users, but is carefully adjusted by sound engineers for CD recording. Linear quantization simplifies mixing of studio signals. The respective number of bits per sample and samples per second approximate signal-to-noise levels and bandwidth typical for phone circuits and high-end recording studios.

Exercise 2 Which steps of the JPEG (DCT baseline) algorithm cause a loss of information? Distinguish between accidental loss due to rounding errors and information that is removed for a purpose.

Example solution:

Accidental loss due to rounding errors:

- RGB → YCrCb conversion
- Discrete cosine transform

Intentional loss:

- resolution reduction in Cr and Cb components
- quantization of DCT results

Exercise 3 How can you rotate by multiples of $\pm 90°$ or mirror a DCT-JPEG compressed image without losing any further information. Why might the resulting JPEG file not have the exact same file length?

Example solution:

Decompress the JPEG file only up to the stage of quantized DCT coefficients. Rotate DCT blocks by negating integer values and/or swapping coefficients (depending on the rotation angle). Then reapply zigzag scan, RLE, DC-coefficient DPCM and Huffman encoding. The results of these lossless steps will differ, which can affect the file length.

<u>Exercise 4</u> Decompress this G3-fax encoded line:
1101011011110111101100110100000000000001

Example solution:

1+14	white	=	110101
3	black	=	10
7	white	=	1111
4	black	=	011
64+63	white	=	11011 00110100
EOL		=	000000000001

⇒ 14 w, 3 b, 7 w, 4 b, 127 w.

<u>Exercise 5</u> You adjust the volume of your 16-bit linearly quantizing soundcard, such that you can just about hear a 1 kHz sine wave with a peak amplitude of 200. What peak amplitude do you expect will a 90 Hz sine wave need to have, to appear equally loud (assuming ideal headphones)?

Example solution:

Refer to the equiloudness diagram on slide 55. The 1 kHz wave is near the hearing threshold, therefore it must have a sound pressure value of near 0 dB_{SPL}. At 90 Hz, the threshold is near 40 dB_{SPL}, therefore the amplitude needs to be 40 dB (= 100× voltage) larger, therefore it should be about 20000.