Correction to exercise 02:

Question 1: Traditionally, for image compression, one would resort to lossy compression. This is due to the fact that the human eye is sensitive only to a certain resolution (transition between two neighboring pixels). Hence, a large compression rate can be obtained without changing the visual quality of an image. Yet, most lossy image compression resort to a lossless compression at the later stages of the algorithm (ex: Huffman and RLE in JPEG).

Question 2:

Each intensity level is coded over 3 bits, hence, there are 8 possible intensity levels.

Question 3:

Entropy of the source:

$$H(I) = \sum_{i \in [0:7]} -P_I(i) \log_2(P_I(i))$$

$$= -\frac{1}{2}\log_2(1/2) - \frac{2}{8}\log_2(1/8) - \frac{3}{16}\log_2\left(\frac{1}{16}\right) - \frac{2}{32}\log_2\left(\frac{1}{32}\right).$$

$$= 1/2 + \frac{3}{4} + \frac{3}{4} + \frac{5}{16}.$$

$$= \frac{8 + 24 + 5}{16} = \frac{37}{16}.$$

$$H(I) = \frac{37}{16} \text{ bits/pixel}$$

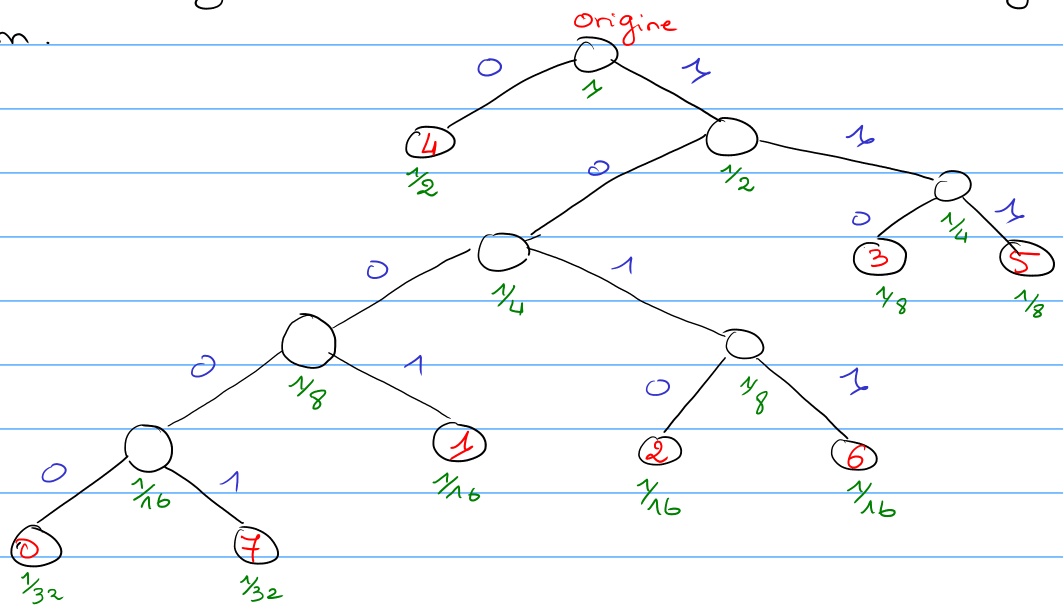
$$\simeq 2,3 \text{ bits/pixel.}$$

Question 4:

In this example, entropy can be measured in bits/pixel. It represents the minimum possible average number of bits necessary to represent the source I.

Question 5:

Let us construct the Huffman tree of this source. There are many possibilities, so we describe here only one of them.



Question 6:

The encoding code book can be obtained from the tree by numbering from the origin down wards any transition to the left by 0 and any transition to the right by 1 (or vice versa). The obtained code is then obtained by:

$0 \longrightarrow 10000$	$4 \longrightarrow 0$
$1 \longrightarrow 1001$	$5 \longrightarrow 111$
$2 \longrightarrow 1010$	$6 \longrightarrow 1011$
$3 \longrightarrow 110$	$7 \longrightarrow 10001$

Question 7:

The intensities with the smaller probabilities have the longest word length, and those with the largest probability (4) have the smallest word length.

Question 8:

The average length of the code is defined by

$$L(c) = \sum_{i \in [0;7]} P_I(i)\, l_c(i)$$

where $l_c(i)$ is the length of the binary codeword coding for i.

Hence: $$L(c) = \frac{1}{2} \cdot 1 + \left(\frac{1}{8} \cdot 3\right) + \left(\frac{1}{8} \cdot 3\right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{16}\right) \cdot 4 + \left(\frac{1}{32} + \frac{1}{32}\right) \cdot 5$$

$$= \frac{37}{16}.$$

Question 9:

This average length is equal to the entropy of the source I. Hence, Huffman code is optimal for this source since

$$L(c) = H(I)$$

This is the case because the source is a 2-adic source, the probability of every intensity is a negative power of 2.