Mittuniversitetet
MID SWEDEN UNIVERSITY

Försättsblad Prov Original

Kurskod	Provkod	Tentamensdatum
DT025A	T101	2018-08-20
Kursnamn	Datateknik AV, Multimediekodning och -distribution	
Provnamn	Tentamen	
Ort	Sundsvall	
Termin	H18	
Ämne	Datateknik	

8982538474

0027927

DT025A

Mittuniversitetet
MID SWEDEN UNIVERSITY

Institutionen för informationsteknologi och medier
Roger Olsson (Roger.Olsson@miun.se)
010 142 8698

2018-08-20

Examination in

Multimedia Coding and Distribution, 7.5 hp

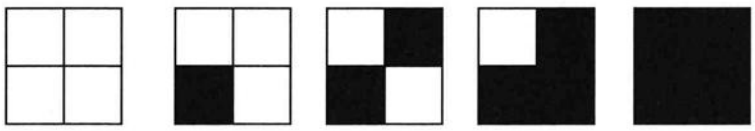
Allowed aids	Dictionary, non-programmable calculator and drawing tools (e.g ruler)
Miscellaneous	Don't forget to: 1. Submit one, or several pages, per solved assignment and NOT several solved assignments per page. 2. Order the handed in pages with the solved assignments in ascending order. 3. Describe and clearly motivate all solutions, assumptions, programming constructs, etc.. 4. Attach all neatly drawn figures that you might want to refer to in a solution. **Failing to follow the above instructions will render point reduction.**

Good Luck!

1. Magoo Inc. recently released a security surveillance application that produces night time photos in the form of 1920 x 1280 pixel gray scale image with 48 gray levels (0=black, 1=dark gray, …, 47=white).

a) A clever engineer at Magoo Inc. came up with the idea to use dithering in order to reduce the required number of bits per pixel to 1 bpp. What is the trade-off that this engineer is willing to make when it comes to the gray scale image properties?

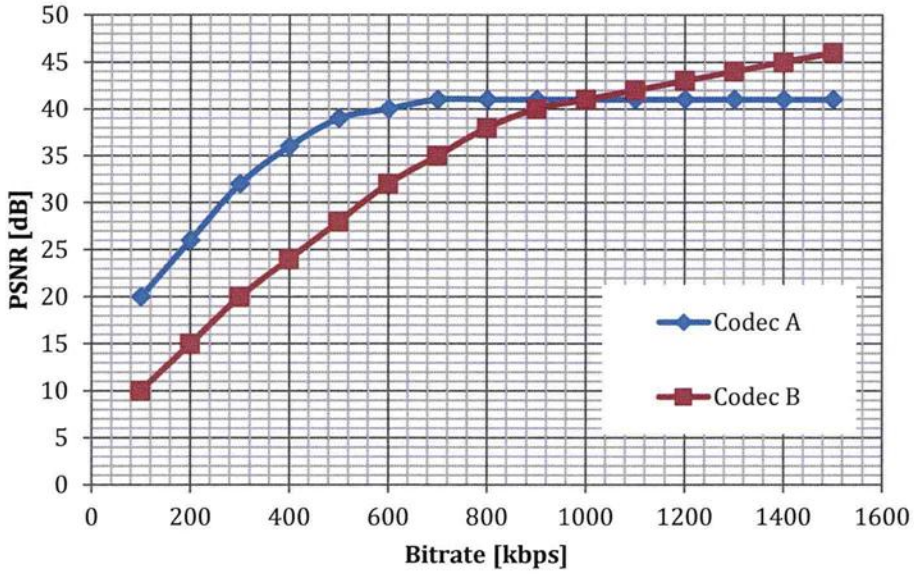
b) What size must the ordered dither square matrix be for the dithering method to be able to represent the full gray scale range of the image? Ordered dither patterns for a 2x2 square matrix is shown below.



Examples of binary dithering patterns

(3 p)

2. The figure below shows rate-distortion curves (RD-graphs) of two encoder/decoders (codec);
a) How much better/worse in dB is codec A compared to codec B at the bitrate of 1.2 Mbps?
b) How much higher/lower bitrate in % is required for codec A compared to codec B at a PSNR quality level of 32 dB?
c) Which codec would preferably be used on a channel that provides a variable bandwidth of 0.7 – 1.0 Mbps?



RD-graph, Codec A and B.

(3 p)

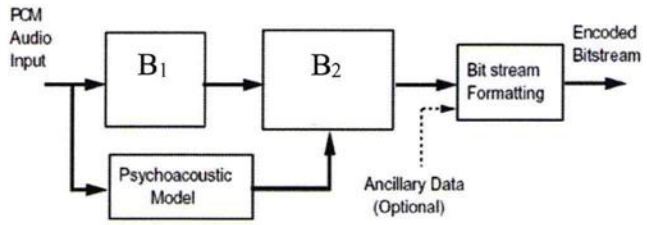
3. A plenoptic camera (PC) enables new image properties and applications such as refocusing, perspective change, and depth extraction in post-processing. The PC images are digital 4D-datasets also known as a lightfield, consisting of a large set of sample vectors **L**(s, t, u, v). It has been determined using measurements that the dynamic range (V_{max}/V_{min}) of each vector component (or sample) s, t, u and v is 56 dB. A first step in compressing this lightfield is to requantize the samples into an as low number of bits as possible. How many bits are required for the signal-to-quantization noise (SQNR) that is introduced by the requantization to be 40 times lower than the minimum signal level?

(4p)

4. Assume that you have a large collection of various images compressed with jpeg. Comment on the feasibility of compressing them into a single video file in terms of compression ratio and ease of access.

(3p)

5. An audio encoder structure is presented in the block diagram of the figure below.



Block diagram of an audio encoder.

The PCM Audio Input in the above figure has the following properties: 5.1 channels ($L_{front}/R_{front}/C_{front}/L_{back}/R_{back}$/LFE). The first five channels use 16 bit per sample, and 48 kHz sampling frequency. The Low Frequency Effects (LFE) or sub-base channel requires only a 4.8 kHz sampling frequency.

a) What is the maximum audio frequency possible in each channel and what PCM bitrate is required for this audio input?
b) The input signal is split and one part is fed to a block containing a psychoacoustic model. How large portion of the encoder's total compression is explicitly caused by the operations in this block? Motivate your answer.
c) Two blocks (B_1 and B_2) are left empty in the block diagram of Figure 1. Name these blocks and explain the operation(s) that is their respective task in the audio compression process.

(5 p)

6. Suppose an alphabet consists of only 3 symbols [A, B, C], and the probability for each of the symbols is P(A) = 0.6, P(B) = 0.3, and P(C) = 0.1. Assume further that both encoder and decoder know that the length of any message is always 5 symbols, which removes the need for a terminator. A message [C, A, B, A, C] is to be transmitted.

a) What is the entropy for the set of symbols used?
b) How many bits are needed to encode the message with fixed bit length?
c) How many bits are needed to encode the message with Huffman coding?
d) How many bits are needed to encode the message arithmetic coding?

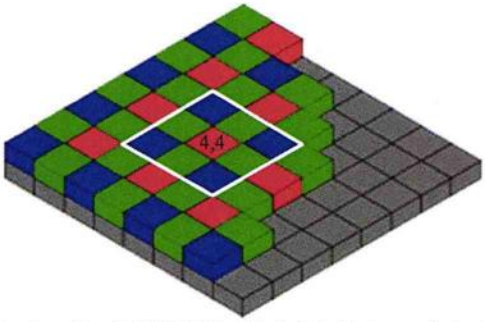
(6 p)

7. When an encoder use inter prediction to reduce temporal redundancy an important operation is motion estimation and -compensation. Performing motion estimation in the form of a sequential search is time consuming.

a) Explain the cause behind why logarithmic search is less computationally complex than sequential search.
b) Compression efficiency is generally enhanced when a motion estimation method adopts half-pixel or quarter-pixel precision. What is the reason behind that?

(4 p)

8. Generally, a conventional camera's image sensor captures only grayscale information. A specific type of color filter or Bayer matrix is applied to the sensor in order to achieve full color images, as shown in the figure below.



Red filter
Green filter
Blue filter

Each pixel's full RGB-triplet is interpolated from the surrounding pixels red, green, and blue information.

a) The debayering/demosaicing implies a tradeoff between color fidelity and spatial resolution. Explain why this tradeoff occurs.

b) A naïve debayering/demosaicing algorithm determines the missing colors of a specific target pixel by averaging over each missing color within a 3x3 pixel region centered at the pixel in question. An example of such a region, centered on pixel P at x=4 and y=4 is shown in the above figure as a white border. The green and blue component of pixel $P(4,4)$ is calculated as the average of the four green, and blue, pixels respectively, i.e.

$$P_G(4,4) = \frac{1}{4}\big(P(4,3) + P(3,4) + P(5,4) + P(4,5)\big)$$

$$P_B(4,4) = \frac{1}{4}\big(P(3,3) + P(5,3) + P(3,5) + P(5,5)\big)$$

Using this algorithm, calculate the full RGB-color (P_R, P_G, P_B) values for the following pixels: P(2,2), P(3,2), P(3,3) using the values presented in the matrix below. Pixels outside the sensor, which might be required for averaging, may be considered to have zero value.

	1	2	3	4
1	**181**	**173**	**30**	**87**
2	**192**	**167**	**127**	**149**
3	**70**	**41**	**245**	**57**

(5 p)

	R	G	B
P(2,2)	167	133	132
P(3,2)	158	127	138
P(3,3)	158	75	245

9. Explain the key properties of the three different frame types that commonly exist in video compression standards, i.e. I-frame, P-frame, B-frame. Draw figures if necessary.

(6p)