北京郵電大學 Queen Mary University of London

For examiners' use only

1	
2	
3	
4	
Total	

EBU5303 A

Joint Programme Examinations 2019/20

EBU5303 Multimedia Fundamentals

Paper A

Time allowed 2 hours

Answer ALL questions

Complete the information below about yourself very carefully.

QM student number										
BUPT student number										
Class number										

Allowed: electronic calculators

INSTRUCTIONS

1. **You must NOT take answer books, used or unused, from the examination room.**
2. Write only with a black or blue pen **and in English**.
3. Do all rough work in the answer book – **do not tear out any pages**.
4. If you use Supplementary Answer Books, tie them to the end of this book.
5. Write clearly and legibly.
6. **Read the instructions on the inside cover.**

Examiners

Dr Marie-Luce Bourguet, Dr Atm Shafiul Alam

Filename: 1920_EBU5303_A No answer book required

Instructions

<u>Before the start of the examination</u>

1) Place your BUPT and QM student cards on the corner of your desk so that your picture is visible.

2) Put all bags, coats and other belongings at the back/front of the room. All small items in your pockets, including wallets, mobile phones and other electronic devices must be **placed in your bag in advance**. **Possession of mobile phones, electronic devices and unauthorised materials is an offence.**

3) Please ensure your mobile phone is switched off and that no alarm will sound during the exam. **A mobile phone causing a disruption is also an assessment offence.**

4) Do not turn over your question paper or begin writing until told to do.

<u>During the examination</u>

1) You must not communicate with or copy from another student.

2) If you require any assistance or wish to leave the examination room for any reason, please raise your hand to attract the attention of the invigilator.

3) If you finish the examination early you may leave, but not in the first 30 minutes or the last 10 minutes.

4) For 2 hour examinations you may **not** leave temporarily.

5) For examinations longer than 2 hours you **may** leave temporarily but not in the first 2 hours or the last 30 minutes.

<u>At the end of the examination</u>

1) You must stop writing immediately – **if you continue writing after being told to stop, that is an assessment offence.**

2) Remain in your seat until you are told you may leave

Question 1

a) This question is about digitisation. Consider a sound wave W with a frequency f = 440 Hz.

[8 marks]

i) What is the sine function representing W?

(1 mark)

ii) What kind of sound is represented by a completely regular sine wave such as W?

(1 mark)

iii) What does the amplitude of W tell us about the sound it represents?

(1 mark)

iv) What is the minimum sampling rate you should use to ensure that you can digitise W without audio aliasing? Justify your answer.

(2 marks)

v) You decide to use 5 bits per sample. How many different values can W take?

(1 mark)

vi) Calculate an approximation of the Signal-to-Quantisation Noise Ratio (SQNR) of W. Explain your calculation.

(2 marks)

	Do not write in this column
Solution:	
(1) $W = \sin(2\pi f t)$, $f = 440$ Hz	
$W(t) = \sin(880\pi t)$	
(2) pure tone	
(3) The amplitude of W tell us the loudness of the	
sound. (How loud the sound is)	
(4) To avoid audio aliasing, according to the	
Nyquist theorem, the sampling rate should larger	
than twice of the sound wave frequency f_m,	
which is $f_s \geq 2f_m = 880$ Hz.	
Hence, the minimum sampling rate is 880 Hz	
(5) $m = 2^n = 2^5 = 32$,	
W can take 32 different values	
(6) SQNR is the ratio of maximum quantization	
level versus maximum quantization noise. In this	
case $SQNR - dB = 20\log_{10}\left(\frac{2^{n-1}}{\frac{1}{2}}\right) = 20\log_{10}(2^n) \approx 6n$	
so $SQNR - dB \approx 6n = 30$.	
	8 marks

b) This question is about colour encoding.

[12 marks]

i) In a true colour image, what is the number of different colours that can be represented? Justify your answer.

(2 marks)

ii) Describe the properties of a fully saturated colour. (brightness)

(2 marks)

iii) In the HSV colour model, how is the grayscale represented?

(2 marks)

iv) What (R, G, B) values would you use to encode an unsaturated dark blue colour?

(2 marks)

v) What (C, M, Y) values would you use to encode a fully saturated bright green colour?

(2 marks)

vi) Yellow ink is spread onto a white sheet of paper. What colour will you see if the paper is illuminated with a blue light? Justify your answer.

(2 marks)

	Do not write in this column
(1) In a true colour image, a 24-bit pixel palette	
will be used, in which 8-bit for each R, G, B value.	
Hence $2^{24} \approx 16.8$ million colours can be represented	
(2) A fully saturated colour ~~is~~ a colour without	
mixture of white	
(3) Assume the largest number for S, V is 100,	
grayscale can be represented as (*, 0, a), $a \in [0,100]$	
intuitively it is the vertical axis of HSV model.	
(4) (20, 20, 100)	
(5) (255, 0, 255)	
(6) Since the yellow ink can absorb its subtractive	
color which is blue, when illuminating with blue	
light. it will absorb all the blue light and	
reflect nothing, so only black can be seen.	
	12 marks

c) Consider a video with the following properties: frame size is 1280-by-720-pixels; colour depth is 24-bits; frame rate is 30 fps; duration is 1 minute.

[5 marks]

i) How much data rate reduction can be achieved by reducing the width and height of this video to half? Prove your answer by calculating the data rates.

(3 marks)

ii) Give two possible strategies for reducing the colour depth of the video to just a third of its original depth.

(2 marks)

	Do not write in this column
(1) Original data rate:	
colour depth: 24-bits/pixel	
bits per frame: 24 × 1280 × 720 = 22118400 bits/frame	
data rate: 22118400 × 30 = 663552000 bps	
= 663.552 Mbps	
Data rate after reduction	
bits per frame: 24 × 640 × 360 = 5529600 bits/frame	
data rate: 5529600 × 30 = 165888000 bps	
= 165.888 Mbps	
Data rate reduction = 497.664 Mbps	
(2) ① converting the video to its grayscale version.	
since the colour depth will be 8 which is a third	
of 24-bit.	
② matching each frame of the video to an	
indexed-based frame with 256-colour CLUT.	
since $2^8 = 256$. colour depth will be 8 as well.	
	5 marks

Question marking: $\frac{}{8} + \frac{}{12} + \frac{}{5} = \frac{}{25}$

Question 2

a) This question is about audio.

[5 marks]

i) Briefly explain what is shown in an audio histogram such as the one in Figure 1. In particular, comment the units used on the X and Y axes.

(3 marks)

ii) Do audio histograms and audio spectrograms represent signals in the same domain? Justify your answer.

(2 marks)

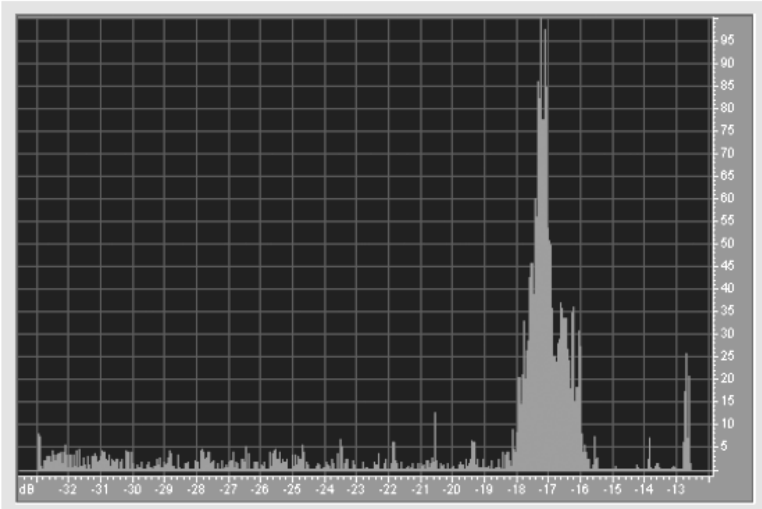


Figure 1: Audio histogram

Do not write in this column

(1) Since audio histogram shows how many samples
at a certain amplitude, X-axis should be amplitude
Y-axis should be numbers of samples. This histogram
shows that the amplitude of this audio concentrated
at −18 to −16 dB which means the audio loudness
don't have too much changes.

(2) They are not in the same domain.
Audio histograms describe relationship between amplitude
and number of samples in time domain.
Audio spectrograms describe how the frequency
spectrum varies with time. so it is in both
time and frequency domain.

5 marks

b) Image lossless compression.

[10 marks]

i) What image property is used in Huffman encoding to achieve compression? Justify your answer.

(3 marks)

ii) Decode the following binary message using the Huffman encoding table provided in Table 1.

E E

Binary message: 1111101111010011011110101101100110111110101011001100100111111

E D E C C B C E C C D D B C E D C C D B C B A

(2 marks)

iii) Assuming that each symbol of Table 1 would normally be encoded using 3 bits (which is enough to encode 5 different symbols), how much compression is achieved in the binary message of question ii) above?

(3 marks)

iv) Consider the following statement: "For compression to remain lossless, an image should be encoded/decoded only once". Is it correct? Justify your answer.

(2 marks)

symbol	probability	code
A	0.40	00
B	0.20	01
C	0.20	10
D	0.10	110
E	0.10	111

Table 1: Huffman encoding table

	Do not write in this column
(1) Statistical distribution, some colours appear more often than other colours. According to this property, we can encode frequent symbols with shorter code, not frequent symbols with longer code to achieve compression.	
(2) EDECCBCECCDDBCEDCCDBCBAEE	
(3) Original size: 25×3 = 75 bits	
Huffman encoding: 11×3+14×2 = 61 bits	

Compression rate = 75 : 61

(4) It is not correct. If the image is encoded in a lossy way, it should only be encoded once since information will loss at this time. However if the encoding method is lossy such as Huffman encoding, no information will be lost in this process so it can be applied for many times.

	10 marks

c) Consider the block diagram shown in Figure 2 and answer the questions below.

[10 marks]

i) Describe what is contained inside the block marked with the label "A".

(2 marks)

ii) What step of the JPEG compression process is marked with the label "B"? Explain this step.

(3 marks)

iii) Explain the use of quantisation tables in the JPEG compression process.

(3 marks)

iv) Which phenomenon of human vision is exploited in chroma sub-sampling?

(2 marks)

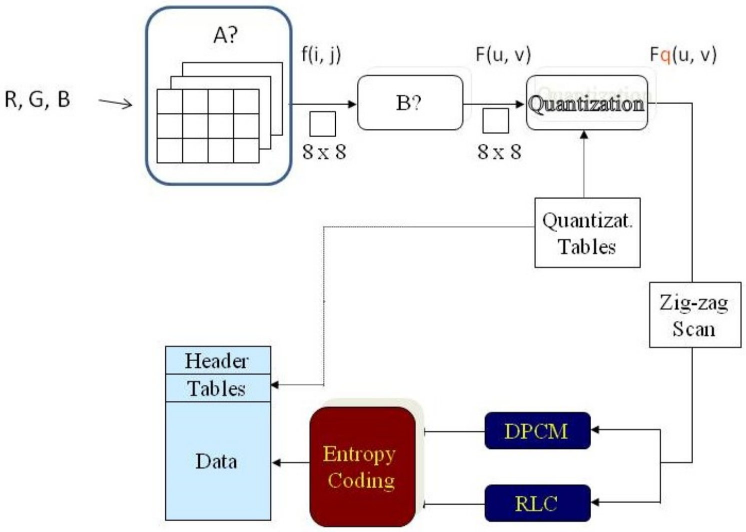


Figure 2: JPEG block diagram

Do not write in this column

(1) In label A, picture is converted from RGB model to a YCbCr model. Then using Chroma subsampling to throw away some chroma information. At last for each 16×16 macroblock outputs 4 8×8 Y block 1 Cb and 1 Cr block.

(2) The label B is DCT transform. In this step, first shift the pixel value by −128. Then applying DCT transform to the 8×8 pixel value matrix, obtain the coefficient for frequency components. The purpose of this step is to seperate out the high frequency components that human eyes are not sensitive to and remove them

(3) The quantisation table is used in quantisation step. It is used to store the quantsize value. In quantisation table, small values are applied to low frequency components which contain fundamental information we don't want to loss, Large values are applied to high frequency components which contain fine details human eyes are hard to captured with.

(4) Human eye is more sensitive to changes in brightness (luminance) than changes in chrominance.

10 marks

Question marking: $\frac{}{5} + \frac{}{10} + \frac{}{10} = \frac{}{25}$

Question 3

a) This question is about MPEG.

[14 marks]

i) What type of MPEG frame makes no prediction?

(1 mark)

ii) What type of MPEG frame is never used as a reference frame during decompression? Why?

(3 marks)

iii) Suppose an MPEG encoder uses the nine-frame sequence IBBPBBPBB. Draw a diagram showing the dependencies between the first 12 frames of a compressed clip produced by this encoder.

(3 marks)

没学

~~iv)~~ Briefly explain the Block Matching Algorithm (BMA) for motion estimation when a maximum motion displacement of x pixels is used.

(4 marks)

~~v)~~ Given a maximum motion displacement of 8 pixels, how many evaluations of the matching criterion are required in the BMA?

(1 mark)

~~vi)~~ Briefly explain the basic principle of fast motion estimation techniques.

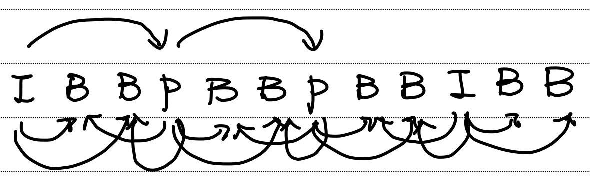
(2 marks)

Do not write in this column

(1) I- frame (Intra-coded frame)

(2) B-frames During decompression, I-frame decompressed without any reference frames. p-frame decompressed with reference frames of previous I-frame or p-frame. B frame decompressed with refrence frames of previous and future I-and p-frames. No frame regard B-frames as reference

(3)



14 marks

b) This question is about perceptual encoding.

[11 marks]

i) With A-law coding, larger signals are represented with greater precision – more data bits – than smaller signals. Is this statement true or false? Justify your answer.

(4 marks)

ii) There are 24 critical bands in the human hearing range, but critical bands for low frequencies are narrower than those for high frequencies. What is this statement telling us about the human ability to distinguish between frequencies?

(3 marks)

iii) What is the threshold of hearing and how does frequency masking affect the threshold of hearing?

(4 marks)

Do not write in this column

(i) This is false. Since human ear is more sensitive to quantisation noise in small signal than in large signals, to presume the dynamic range. small signals should be represented with greater precision – more data bits – than larger signals.

(2) Human ability to distinguish frequencies decreases nonlinearly from low to high frequencies. Since human ear can distinguish better in lower frequencies. the critical bands for lower frequencies is narrower.

(3) Threshold of hearing: the minimum level of the sound can be heard.
The frequency masking will raised the threshold of hearing within a critical band in the presence of masking tone

	11 marks

Question marking: $\frac{}{14} + \frac{}{11} = \frac{}{25}$

Question 4

a) This question is about MP3.

[15 marks]

i) In MP3, one way to reduce the amount of data in the compressed signal is to use scaling factors that increase the quantisation error where it doesn't matter. Briefly explain how the parts of the signal that will be multiplied by a large scaling factor can be found.

(5 marks)

ii) Say that an uncompressed band value is 10,000 and values from all bands are quantised by dividing by 128 and rounding down. What is the quantisation error? Show your calculations.

compander (rounding down)

(3 marks)

iii) Now suppose that this band requires less precision because of a strong masking tone, and that it should be scaled by a factor of 0.1. Recalculate the quantisation error.

(3 marks)

iv) With an MP3 bitrate of 128 kbit/s, calculate the compression ratio that is achieved on a CD quality digital audio signal (CD quality = 44100 samples per second, stereo and 16 bits per channel).

(2 marks)

v) What is meant by "Average Bit Rate" (ABR)?

(2 marks)

Do not write in this column

(1) Before the quantisation and scaling step in mp3, at the psychoacoustic analysis, it will identify the masking and masked tones and outputs a set of SMR values. With the lower SMR, it can be multiplied by a large scaling factors, same for these parts. the quantisation error will be lower than the masking threshold, i.e. quantisation errors disappear less to this parts

(2) $\left\lfloor \frac{10000}{128} \right\rfloor = 78$, error: $10000 - 78 \times 128 = 16$

error rate: $\frac{16}{10000} = 0.0016$

(3) $\left\lfloor \frac{10000 \times 0.1}{128} \right\rfloor = 7$, error: $10000 - \frac{7 \times 128}{0.1} = 1040$

error rate: $\frac{1040}{10000} = 0.104$

(4) sampling rate = 44.1 k sample / second
quantisation = 16 bits / sample stereo : 2
Bit rate = 44.1 K × 16 × 2
= 1411200 bps = 1411.2 kbps
Compression ratio = $\frac{1411.2 \text{ Kbps}}{128 \text{ Kbps}}$
= 441 : 40

(5) Average Bit Rate is a type of VBR (Variable bit rate). The bit rate is variable in order to achieve a consistent quality, besides. the bit rate is bounded to an average value to achieve a predictable file size chosen by user.

	15 marks

b) This question is about DVB-S.

[10 marks]

i) What compression standard is used for source coding in DVB-S?

(2 marks)

ii) What is the purpose of energy dispersal?

(2 marks)

iii) How does the Reed-Solomon Error Protection scheme work?

(2 marks)

iv) Assuming a symbol rate of 27.5 MS/s, QPSK modulation, Reed-Solomon code with rate (204, 188), and a code rate of ¾ are used, calculate the bit stream net data rate. Show your calculations.

(4 marks)

Do not write in this column

(1) MPEG-2

(2) The purpose is to achieve a power-density spectrum of a modulated signal as even as possible

(3) Reed-Solomon Protection schemes correct bits at decoder by adding some "redundant bit". Specifically, assume there are k data in a packet. we add some bit to obtain n data. Then the decoder can correct t symbols that contain errors ($2t = n - k$)

(4) QPSK modulation means 2 bits/symbol.
symbol rate = 27.5 MS/s
gross_datarate = 55 Mb/s
net_datarate (RS) = $55 \times \frac{188}{204} \approx 50.69$ Mb/s
net_datarate = net_datarate (RS) × code rate
$= \frac{3}{4} \times 50.69$
≈ 38.01 Mb/s

	10 marks

Question marking: $\frac{}{15} + \frac{}{10} = \frac{}{25}$

Do not write in this column

Do not write in this column

Do not write in this column