

NANYANG TECHNOLOGICAL UNIVERSITY**SEMESTER 2 EXAMINATION 2021-2022****EE6403 – DISTRIBUTED MULTIMEDIA SYSTEMS**

April / May 2022

Time Allowed: 3 hours

INSTRUCTIONS

1. This paper contains 5 questions and comprises 5 pages.
 2. Answer all 5 questions.
 3. All questions carry equal marks.
 4. This is a closed book examination.
 5. Unless specifically stated, all symbols have their usual meanings.
-

1. In an application, Karhunen-Loeve Transform (KLT) is used to compress a class of images. The images are partitioned into 2×2 pixel blocks, and ordered lexicographically row-by-row to form 4×1 vectors. The covariance matrix of the vectors is given by

$$\mathbf{C} = \begin{bmatrix} 2.1503 & 0.9073 & 3.3534 & 1.7945 \\ 0.9073 & 1.8596 & -0.0251 & 3.6291 \\ 3.3534 & -0.0251 & 6.7519 & -0.0401 \\ 1.7945 & 3.6291 & -0.0401 & 7.3082 \end{bmatrix}.$$

The covariance matrix \mathbf{C} has four eigenvalues, $\lambda_1, \lambda_2, \lambda_3$ and λ_4 , with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and \mathbf{v}_4 .

- (a) Find the eigenvalues λ_1 and λ_2 given that

$$\mathbf{v}_1 = \begin{bmatrix} 0.3780 \\ 0.3780 \\ 0.3780 \\ 0.7559 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -0.2965 \\ 0.2224 \\ -0.8154 \\ 0.4448 \end{bmatrix}.$$

(4 Marks)

Note: Question No. 1 continues on page 2.

- (b) Given that $\lambda_3 = 0.05$, $\lambda_4 = 0.02$, $\mathbf{v}_3 = [0.32 \ 0.80 \ -0.16 \ m]^T$, and $\mathbf{v}_4 = [n \ -0.41 \ -0.41 \ 0]^T$, where m and n are real numbers, find the values m and n . (4 Marks)

- (c) A pixel block \mathbf{I} after zero-mean centering is given by

$$\mathbf{I} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

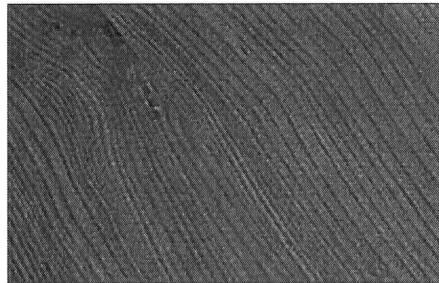
The pixel block is to be compressed using KLT with the covariance matrix \mathbf{C} . A user chooses to use only the most dominant 1 principal component to compress pixel block \mathbf{I} . Calculate the total square error between the original pixel block and the reconstructed pixel block.

(7 Marks)

- (d) KLT is used to compress a new class of images. The new images are partitioned into 2×2 pixel blocks, and ordered lexicographically row-by-row to form new 4×1 vectors. The covariance matrix of these new vectors is given by \mathbf{D} . Assume that the covariance matrix \mathbf{D} has eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and \mathbf{u}_4 , with corresponding eigenvalues of 4, 3, 2 and 1. Express and simplify the covariance matrix \mathbf{D} in terms of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ and \mathbf{u}_4 .

(5 Marks)

2. (a) Draw an encoder diagram for predictive coding in image compression. Clearly label all the key components. Briefly discuss the effectiveness of predictive coding in compressing the image given in Figure 1.

**Figure 1**

(6 Marks)

Note: Question No. 2 continues on page 3.

- (b) In an image compression scheme similar to the baseline JPEG, a user partitions an image into 4×4 image pixel blocks, performs 4×4 two-dimensional Discrete Cosine Transform (2-D DCT) for each block, and quantizes the transform coefficients using 4×4 quantization table \mathbf{Q} given by

$$\mathbf{Q} = \begin{bmatrix} 20 & 30 & 40 & 50 \\ 30 & 40 & 50 & 60 \\ 40 & 50 & 60 & 70 \\ 50 & 60 & 70 & 80 \end{bmatrix}.$$

- (i) After the quantization step, the quantized coefficients go through the entropy encoding. List the key step(s) in the entropy encoding to perform image compression.
- (ii) The user finds that the quality of the decompressed images using the quantization table is lower than expected as they contain some distortion. The user would like to improve the quality of the decompressed images. Suggest a suitable new quantization table to achieve the goal, and justify your answer briefly.
- (iii) The scheme is used to compress two 4×4 image pixel blocks \mathbf{A} and \mathbf{B} :

$$\mathbf{A} = \begin{bmatrix} 20 & 10 & 20 & 10 \\ 10 & 20 & 10 & 20 \\ 20 & 10 & 20 & 10 \\ 10 & 20 & 10 & 20 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 20 & 20 & 20 & 20 \\ 20 & 20 & 20 & 20 \\ 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 \end{bmatrix}.$$

State which image pixel block will likely experience more distortion in the decompressed image after image decompression. Justify your answer briefly.

(14 Marks)

3. (a) In an application, a user would like to compress a video using the MPEG-2 standard. The compressed MPEG-2 video has a Group-of-Picture (GOP) structure of IBBBPBBPB, and a frame rate of 30 frames per second (fps). The average compression ratios for the I-frames, P-frames, and B-frames of this video are assumed to be $10:1$, $p:1$, and $q:1$, respectively, where p and q are real numbers. The video is to be displayed using the following format:

Resolution of the luminance plane: 720×480 pixels

Color depth of every pixel in each luminance and chrominance plane: 8 bits/pixel

Chroma subsampling format: 4:2:2

Note: Question No. 3 continues on page 4.

The target average bitrate of the compressed video is to be less than 3.1×10^6 bits/s. Assume that $q=2p$, find the conditions for the p and q values that can satisfy the bitrate requirement. State any assumption(s) you make in the calculation. Comment briefly whether the conditions for the p and q values are reasonable in practical applications.

(10 Marks)

- (b) State an advantage and a disadvantage of three step search when compared with full search in motion estimation.

(4 Marks)

- (c) With the aid of a diagram, explain briefly how a Frequency Division Multiplexing (FDM) transmitter performs multiplexing for signals from different users.

(6 Marks)

4. (a) A Cyclic Redundancy Check (CRC) system uses a generator polynomial of $P(X) = X^3 + aX + 1$, where a is a binary number. In a particular transmission, a codeword polynomial of $R(X) = X^6 + X^4 + X^3$ is received at the receiver. It is known that transmission errors occur during this transmission, and the CRC at the receiver can detect the transmission errors.

- (i) Find the value a .
- (ii) Calculate the remainder obtained using the CRC at the receiver.
- (iii) Can the original dataword at the transmitter be recovered by the CRC? Justify your answer briefly.

(9 Marks)

- (b) List 3 key advantages of Transmission Control Protocol (TCP) over User Datagram Protocol (UDP).

(5 Marks)

- (c) Briefly discuss the objective of Address Resolution Protocol (ARP) and its operation in the Ethernet.

(6 Marks)

5. (a) Write down the order from the largest to the smallest the number of trainable parameters / weights for: (i) a convolutional layer, (ii) a max pooling layer, and (iii) a fully connected layer of a typical Convolutional Neural Network (CNN). Justify your answer briefly.

(4 Marks)

Note: Question No. 5 continues on page 5.

- (b) In an image classification application using CNN, the output scores from the last fully connected layer of a training sample are given by $[0.2, 0.1, 0.4, 0.3, 0.8]$. Given that the ground truth output labels of this training sample are given by $[0, 0, 0, 0, 1]$, calculate the softmax loss for this training sample. (6 Marks)
- (c) With the aid of a diagram, briefly describe the key steps how a Vision Transformer (ViT) can perform image classification. List two advantages of Vision Transformer over Convolutional Neural Networks in performing image classification. (10 Marks)

END OF PAPER

EE6403 DISTRIBUTED MULTIMEDIA SYSTEMS

Please read the following instructions carefully:

- 1. Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**
2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.
3. Please write your Matriculation Number on the front of the answer book.
4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.