MACHAKOS UNIVERSITY

University Examinations for 2023/2024 Academic Year

SCHOOL OF ENGINEERING AND TECHNOLOGY

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF SCIENCE (TELECOMMUNICATION AND INFORMATION TECHNOLOGY)

SCS 474: COMPRESSION TECHNIQUES

DATE: **TIME:**

INSTRUCTIONS

This paper consists of **FIVE** questions. Answer Question **ON**E and any other **TWO** questions

QUESTION ONE (COMPULSORY) (30 MARKS)

a) State **two** needs for the compression of data (2 marks)

b) (i) For the given data sequence, write down the compressed version using the run-length encoding scheme.
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1

(ii) Develop a pseudo code for the run-length encoding algorithm (6 marks)

c) The code in the appendix (last page of the paper) implements a run-length encoding scheme.
(i) Comment it out;
(ii) Provide a dry-run output. (4 marks)

d) Distinguish between **lossy** and **lossless** data compression. (2 marks)

e) State the Nyquist's Sampling Theorem. (1 mark)

d) For each of the following media types, briefly discuss how Nyquist's Sampling Theorem affects the quality of the data and the form in which sampling effects manifest themselves in the actual data.
i. graphics,
ii. images,
iii. audio
iv. video, (4 marks)

e) Calculate the uncompressed digital output, i.e. data rate, if a video signal is sampled using the following values:

- 25 frames per second
- 160 x 120 pixels
- True (Full) colour depth

(4 marks)

f) A discrete source generates three independent symbols A, B and C with probabilities

0.9, 0.08 and 0.02 respectively. Determine the entropy of the source.

(3 marks)

g) Design binary Huffman code for a discrete source of three independent symbols A, B and C with probabilities 0.9, 0.08 and 0.02 respectively. Determine the average code length for the code.

(4 marks)

QUESTION TWO (20 MARKS)

a) Define each of the following with respect to codes

i. Prefix codes

ii. Code efficiency

(2 marks)

b) A source emits symbols {A, B, C, D,E} with corresponding probabilities given in Table 1.

Table 1

Source symbol	A	B	C	D	E
Probability	$^{8}/_{50}$	$^{9}/_{50}$	$^{7}/_{50}$	$^{20}/_{50}$	$^{6}/_{50}$

i. Estimate the Shannon entropy of this source

(2 marks)

ii. Derive the Shannon-Fano code for each symbol

(5 marks)

iii. Derive the Huffman code for each symbol

(5marks)

iv. Estimate the number of ones and zeros required to encode symbols from the above source using the Shannon-Fano and Huffman Codes derived.

(2 marks)

QUESTION THREE (20 MARKS)

a) Given the following string as input:
VETITIVE
with the initial dictionary below, encode the string with the LZW algorithm showing the intermediate steps.

Index	Entry
1	E
2	I
3	T
4	V

(8 marks)

b) One simple approach to video compression is to compress *each frame* of the video using the JPEG pipeline as done in motion JPEG (M-JPEG) video format. What disadvantages of this simple approach make it an unpopular choice as a video compression standard?

(2 marks)

c) (i) Explain the key differences between I-frames, P-frames and B-frames in MPEG-2 video compression.
(ii) Describe the advantages and disadvantages of using B frames.
(iii) Which of the above kind of frames typically requires the most and the least number of bits to encode?
(v) What is the reason for inserting I-frames into the compressed output stream relatively frequently?

(5 marks)

QUESTION FOUR (20 MARKS)

a) What is the main difference between the H.261 and MPEG video compression algorithms?

(2 marks)

b) MPEG has a variety of different standards, i.e. MPEG-1, MPEG-2, MPEG-4, MPEG-7 and MPEG-21. Why have such standards evolved? Give an example target application for each variant of the MPEG standard.

(7 marks)

c) Briefly describe the four basic types of data redundancy that data compression algorithms can apply to audio, image and video signals.

(6 marks)

QUESTION FIVE (20 MARKS)

a) Encode the following steam of characters using decimal arithmetic coding compression:

MEDIA

You may assume that characters occur with probabilities of M = 0.1, E = 0.3, D = 0.3, I = 0.2 and A = 0.1.

(12 marks)

b) Show how your solution to (c) would be decoded.

(3 marks)

Appendix

```
import java.io.*;
public class GFG {
   public static void encoding(String str)
   {
      int n = str.length();
      for (int i = 0; i < n; i++) {

         // (Provide comment for this section)
         int count = 1;
         while (i < n - 1
               && str.charAt(i) == str.charAt(i + 1)) {
            count++;
            i++;
         }
         // Provide comment for this section
         System.out.print(str.charAt(i) + "" + count);        }
   }

   // Provide comment for this section
   public static void main(String[] args)
   {
      String str = "wwwwaaadexxxxxxywww";
      encoding(str);
   }
}
```