COMP5422 Multimedia Computing

Midterm Test

(This is a closed book, closed note test.)

Name:______________________ Student number:________________

Part A. True / False (12 points, 2 points each)

1. The HSI color model is suitable for image acquisition. **[F]**
2. The hue represents the dominant color of an object. **[T]**
3. The binary representation of 12 is “1100”. **[T]**
4. The coding tree for an input symbol stream is unique by using the Shannon-Fano algorithm. **[F]**
5. The run-length coding is one of the simplest entropy coding methods for data compression. **[F]**
6. In variable length coding, symbols that occur more frequently will have longer code-words than symbols that occur less frequently. **[F]**

Part B. Multiple choices. (15 points; 3 points each)

The following questions have <u>at least one</u> correct answer.

1. Achromatic light can be described by **[A]**

   A. Intensity; B. Radiance; C. Luminance; D. Brightness.

2. If we use 9 bits to represent each pixel in a grey level image, what is the grey level range of this image? **[F]**

   A. [0, 255]; B. [0, 1024]; C. [1, 256];
   D. [0, 999]; E. [0, 512]; F. [0, 511].

3. Suppose that “X” and “Y” are two symbols in a message. If “X” is coded as “101”, then “Y” cannot be coded as **[B, E]**

   A. 01; B. 10; C. 110; D 011; E 1011.

4. Suppose the minimum and maximum frequencies of an audio signal are 10K HZ and 50K HZ. To preserve all the content of this signal, the sampling rate should be at least **[E]**

   A. 20K HZ; B. 40K HZ; C. 100K HZ;
   D. 60K HZ; E. 80K HZ; F. 120K HZ.

5. Which one(s) of the following audio coding schemes does(do) not exploit the signal consistency over time? **[B]**

   A. LPC; B. PCM; C. DPCM; D. ADPCM;

Part C. Short-answer questions. (38 points)

1. Suppose the true frequency of the audio signal is 20K HZ but the sampling frequency is 36K HZ. Is the sampled signal aliased? If yes, what is the alias frequency? (3 points)

   Answer:

   Yes. (1 point)
   The aliased frequency is 16K HZ**.** (2 points)

2. In audio coding, why do we need non-uniform quantization and what are the two commonly used non-uniform mapping functions? (5 points)

   Answer:

   Human auditory system exhibits a logarithmic sensitivity. It is more sensitive at small amplitude range and less sensitive at large-amplitude range. Thus we can assign smaller quantization step-size for small-amplitude regions and larger quantization step-size for large-amplitude regions. (3 points)

   The two commonly used nonlinear mapping functions are A-law and Mu-law. (2 point)

3. Using the primary R, G and B colors, can we reproduce all the visible spectrum colors? Why? (3 points)

   Answer:

   No. (1 point)
   The R, G, and B primary color can only reproduce the colors within the triangle determined by them in the CIE chromaticity diagram. Clearly they cannot cover all the visible colors. (2 points)

4. The amounts of red, green and blue for a particular color are called tri-stimulus X, Y and Z. In the CIE chromaticity diagram, how the color is specified? (3 points)

Answer: (1 point each)

x= X/(X+Y+Z)
y= Y/(X+Y+Z)
z= 1-x-y

5. Please code string "HEELLOOO" using run length coding. (3 points)

Answer:

H1E2L2O3.

6. Suppose the original signal is X = [1.1, 0.8, 0.7, 1.3]. After quantization, the signal is quantized to Y = [1, 1, 1, 1]. What is the SQNR of Y? (5 points)

Answer:

N=X-Y = [0.1 -0.2-0.3 0.3]; (2 points)
The SQNR is 12.44dB. (3 points)

7. What is the 6dB rule? (4 points)

Answer:
Suppose we choose a quantization accuracy of N bits per sample with one bit is used to indicate the sign of the sample. Then the maximum signal value is mapped to $2^{N-1}-1$ ($\approx 2^{N-1}$) and the most negative signal is mapped to -2^{N-1}. Suppose the uniform quantization interval is 1, so the quantization error is at most ½, the half of the interval. (2 points)

In the worst case that the noise is the maximum ½, the SQNR can be simply expressed:

$$SQNR = 20\log_{10}\frac{V_{signal}}{V_{quan\_noise}} = 20\log_{10}\frac{2^{N-1}}{\frac{1}{2}}$$

$$= 20 \times N \times \log 2 = 6.02\,N(\mathrm{dB})$$

We can see that for a uniformly quantized source, adding 1 bit/sample can improve the SNR by 6dB. This is called the 6dB rule. (2 points)

8. Suppose we have an image

69	153
28	90

What is the 1st bit plane of this image? (4 points)

Answer:

0	0
0	1

9. Please explain why and what are **safe RGB** colors? (4 points)

Answer:

In many applications, it makes no sense to use more than a few hundred colors. A subset of colors that are likely to be reproduced reasonably and independently of viewer hardware is called safe RGB color. (2 points)

216colors have become the de facto standard for safe colors. Each of 216 safe colors is formed from three RGB values but each value can only be one of the six possible values (0, 51, 102, 153, 204 or 255). Thus there are totally 6^3=216 possible values. (2 points)

10. After digitization, a grey level image has 400 rows and 200 columns, and each pixel of it is stored by 10 bits. (a) What is the **spatial resolution** of it? (b)What is the **grey-level resolution** of it? (c) How many **bytes** do we need to store this image? (4 points)

Answer:

(a) 400*200; (1 points)
(b) 1024; (1 points)
(c) 400*200*10/8 = 100,000 bytes. (2 points)

Part D. Long-answer questions. (35 points)

1. Suppose we have an 8×8 image as follows:

99	99	99	99	99	99	99	99
20	20	20	20	20	20	20	20
0	0	0	0	0	0	0	0
0	0	50	50	50	50	0	0
0	0	50	50	50	50	0	0
0	0	50	50	50	50	0	0
0	0	50	50	50	50	0	0
0	0	0	0	0	0	0	0

(a) What is the entropy of this image? (4 points)
(b) Construct the Huffman tree to encode the values of this image and show the resulting code for each intensity value. (4 points)
(c) What is the average number of bits needed for each pixel using your Huffman code? (2 points)

Answer:

(a)

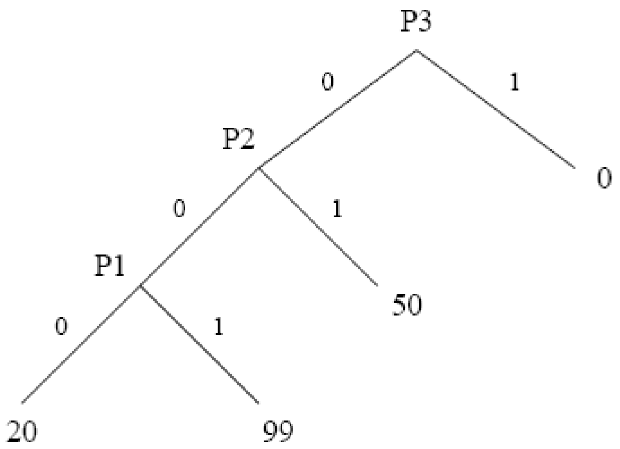
$$P_{20} = P_{99} = 1/8,\ P_{50} = 1/4,\ P_0 = 1/2.$$

(2 points)

$$\eta = 2 \times \frac{1}{8}\log_2 8 + \frac{1}{4}\log_2 4 + \frac{1}{2}\log_2 2 = \frac{3}{4} + \frac{1}{2} + \frac{1}{2} = 1.75$$

(2 points)

(b)



(3 points)

Resulting code: 0: “1”, 50: “01”, 20: “000”, 99: “001” (3 points)
or: 0: “1”, 50: “01”, 99: “000”, 20: “001”
or: 0: “0”, 50: “10”, 20: “111”, 99: “110”
or: 0: “0”, 50: “10”, 99: “111”, 20: “110”

(c) Average number of bits = 0.5×1 + 0.25×2 + 2×0.125× 3 = 1.75. (2 points)

2. Suppose we have the symbol stream “ABCB” to code. For simplicity let’s assume that both the encoder and decoder know that the length of the message is 4 so that there is no need for a terminator added to the symbol stream. Please use arithmetic coding to code the stream. (14 points)

Answer:

(a) Start by assigning each symbol to the probability range 0 –1. Sort symbols with the highest probability first.

Symbol	Range
B	[0.0,0.5)
A	[0.5, 0.75)
C	[0.75,1.0)

(1 point)

1. The range of first symbol A is [0.5, 0.75), so low = 0.5, high = 0.75 and range = high – low = 0.25. (1 point)

2. The next symbol being coded is B. Thus high_range = 0.5 and low_range = 0.0. Now:

high = 0.5 + 0.25*0.5 = 0.625
low = 0.5 + 0.25*0 = 0.5

So, the range of stream AB is [0.5, 0.625). (2 points)

3.For stream AB, low = 0.5, high = 0.625 and range = 0.625 – 0.5 = 0.125. The next symbol being coded is C. Thus, high_range = 1.0 and low_range = 0.75. Now:

high = 0.5 + 0.125*1.0 = 0.625;

low = 0.5 + 0.125*0.75 = 0.59375.

So, the range of stream ABC is [0.59375, 0.625). (2 points)

4.For ABC, low = 0.59375, high = 0.625 and range = 0.625-0.59375 = 0.03125. The next symbol being coded is B. Thus, high_range = 0.5, low_range = 0. Now:

high = 0.59375+0.03125*0.5 = 0.609375

low = 0.59375+0.03125*0 = 0.59375

So, the range of stream ABCB is [0.59375, 0.609375). (2 points)

(b) Next we compute the binary codeword.

If we assign 1 to the first binary fraction bit, i.e. 0.1 binary, its decimal value(code) = value(0.1) = 0.5 decimal < low = 0.59375. Then we go to the next loop. (1 point)
If we assign 1 to the second binary fraction bit, i.e. 0.11 binary, its decimal value(code) = value(0.11) = 0.75 > high = 0.609375. Then we let the second bit be 0. Since value(0.10) = 0.5 < low = 0.59375, we continue. (1 point)
If we assign 1 to the third binary fraction bit, i.e. 0.101 binary, its decimal value(code) = value(0.101) = 0.625 > high = 0.609375. Then we let the third bit be 0. Since value(0.100) = 0.5 < low = 0.59375, we continue. (1 point)
If we assign 1 to the fourth binary fraction bit, i.e. 0.1001 binary, its decimal value(code) = value(0.1001) = 0.5625 < low = 0.59375. Then we let the fourth bit be 1. Since value(0.1001) = 0. 5625 < low = 0.59375, we continue. (1 point)
If we assign 1 to the fifth binary fraction bit, i.e. 0.10011 binary, its decimal value(code) = value(0.10011) = 0.59375 =low = 0.59375< high. (1 point)
Then we stop and get the binary codeword 0.10011. (1 point)

3. Suppose we wish to code the sequence [f_1, f_2, f_3, f_4, f_5] = [100, 98, 101, 105, 103] using lossless predictive coding (LPC). The predictor we use is

$$\hat{f}_n = \left\lfloor \tfrac{1}{2} f_{n-1} + \tfrac{1}{2} f_{n-2} \right\rfloor$$

$$e_n = f_n - \hat{f}_n$$

(a) Please calculate the signal we need to transmit in the encoder side. (4 points)
(b) Please draw the diagram of the decoder of LPC. (3 points)
(c) Please reconstruct the signal in the decoder side. (4 points)

Answer:

(a) Let $f_0 = f_1 = 100$.

$\hat{f}_2 = \left\lfloor \tfrac{1}{2} f_1 + \tfrac{1}{2} f_0 \right\rfloor = 100;\ e_2 = 98 - 100 = -2$ (1 point)

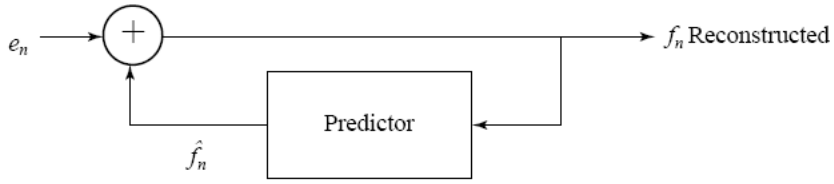
$\hat{f}_3 = \left\lfloor \tfrac{1}{2} f_2 + \tfrac{1}{2} f_1 \right\rfloor = 99;\ e_3 = 101 - 99 = 2$ (1 point)

$\hat{f}_4 = \left\lfloor \tfrac{1}{2} f_3 + \tfrac{1}{2} f_2 \right\rfloor = 99;\ e_4 = 105 - 99 = 6$

$\hat{f}_5 = \left\lfloor \tfrac{1}{2} f_4 + \tfrac{1}{2} f_3 \right\rfloor = 103;\ e_5 = 103 - 103 = 0$ (1 point)

The encoder will transmit [f_1, e_2, e_3, e_4, e_5] = [100, -2, 2, 6, 0]. (1 point)

(b)



(3 points)

(c) Now the decoder receives $[f_1, e_2, e_3, e_4, e_5] = [100, -2, 2, 6, 0]$.
Let $f_0 = f_1 = 100$.

$\hat{f}_2 = \left\lfloor \frac{1}{2} f_1 + \frac{1}{2} f_0 \right\rfloor = 100; \quad f_2 = \hat{f}_2 + e_2 = 100 - 2 = 98$ (1 point)

$\hat{f}_3 = \left\lfloor \frac{1}{2} f_2 + \frac{1}{2} f_1 \right\rfloor = 99; \quad f_3 = \hat{f}_3 + e_3 = 99 + 2 = 101$ (1 point)

$\hat{f}_4 = \left\lfloor \frac{1}{2} f_3 + \frac{1}{2} f_2 \right\rfloor = 99; \quad f_4 = \hat{f}_4 + e_4 = 99 + 6 = 105$ (1 point)

$\hat{f}_5 = \left\lfloor \frac{1}{2} f_4 + \frac{1}{2} f_3 \right\rfloor = 103; \quad f_5 = \hat{f}_5 + e_5 = 103 + 0 = 103$ (1 point)

Finally, the decoder reconstructs perfectly the original signal.