1102 Video Compression - Midterm Exam (Total score: 100)

a d e 1. (36%) Multiple-choice Questions

1.1. (14%) For video compression using motion compensated predictive coding, which the following statement(s) is (are) true?

(a) Using a large block-size motion estimate has less overhead for sending motion vectors than a small block-size ME. fewer blocks

(b) The results of block matching using different matching criteria will always be different. ?

(c) Using a large block-size motion estimate results in lower prediction error in inter-coding when a large block involves more than one moving object. smaller

(d) One can use the mean square error to measure similiarity in a block matching algorithm.

(e) Asymmetric motion block shapes can better fit moving objects, leading to smaller prediction errors. ?

(f) Motion vectors are all we need to recover the blocks in motion compensation in the decoding process. also needs residual (last frame)

(g) Enabling half-pixel motion estimation can always lead to higher objective quality for video compression.

d e 1.2. (12%) Which of these codes can be Huffman for any probability assignment? prefix-free

(a) {0, 10, 01, 11}

(b) {00, 01, 10, 110}

(c) {01, 10}

(d) {0, 10, 11}

(e) {1, 01, 001, 000}

(f) {1, 01, 11}

draw the tree

b c e 1.3. (10%) which the following statement(s) is (are) true?

(a) The HVS is more sensitive to chrominance details. luminance

(b) The higher the frame rate, the more smoothly a video plays.

(c) Removing high frequency information, such as edges, object texture, is helpful and good to compression.

(d) HVS is more sensitive to high contrast. good for compression

(e) The statement, "I saw a superman on campus today," bears much information. surprising

2. (32%) Answer the following questions with a brief explanation.

2.1. (9%) Talking about redundancy, please name at least three types of redundancy and describe their definition.

2.2. (8%) How many bits do we need to store a 1-minute 30 frame-per-second qcif uncompressed video in the RGB color space? (Note that qcif = 176*144, a pixel in a video frame consists three color components, each of which takes 8 bits. **please show how to calculate it and do not need to evaluate it**)

2.3. (4%) Following the previous question (2.2.), if we transform the video into the YCbCr 4:2:0 format, how many bits do we need then?

2.4. (6%) Please explain why using PSNR may be irreliable to measure the quality of decoded videos? To address the irreliablilty of PSNR, what measure can we also use to assess video quality?

2.5. (5%) Please decoding code 000101 with m=5 using Golomb coding.

3. (12%) Motion Estimation – for a search range of 15×15 pixels
   (a) (6%) How many search comparisons are required for Full Search and Three-step Search?
   (b) (6%) For one large-pattern searches and one small-pattern search, how many searches are required for Diamond Search and Hexagon Search?

4. (20%, 4% for each question) Please briefly describe the following functional units for JPEG encoding.
   (a) Level shift; (b) DCT; (c) Quantization; (d) ZigZag Scan; (e) DC prediction

2.1.

redundancy
- spatial : neighboring pixels are similar
- temporal : similarities b/w successive frames
- coding : using more bits than necessary
- irrelevant : info. not perceptually important to eyes

2.2. frame resolution × color channel × bits/channel × frame rate × duration

$$\left(\frac{\text{pixel}}{\text{frame}} \times \frac{\text{channel}}{\text{pixel}} \times \frac{\text{bit}}{\text{channel}} \times \frac{\text{frame}}{\text{s}} \times \text{s} = \text{bit}\right)$$

$$= (176 \times 144) \times 3 \times 8 \times 30 \times 60$$

2.3. Y : 176×144

Cb, Cr : $\frac{176}{2} \times \frac{144}{2}$

2.4. 1. PSNR measures pixel-wise difference and may not reflect perceptual quality.

2. DSCQS (double stimulus continuous quality

scale)

2.5

000 | | 0 |

3 ⋮ q

1 ⋮ r

$n = 5 \Rightarrow b = \lceil \log_2 5 \rceil = 3$

$\Rightarrow 2^b - m = 2^3 - 5 = 3$

$X = m \cdot q + r$

$= 5 \cdot 3 + 1$

$= 16$ #

0	0	
1	1	✓
2	2	
3	6	
4	8	

4.