UNIVERSITY of CALIFORNIA
Santa Barbara
Department of Electrical and Computer Engineering

ECE 178 **Sample Problem Solutions**

B.S. Manjunath
TAs: J. Tešić, M. Zuliani

Question 1

1. Using homomorphic filtering:
   $s(x,y) = f(x,y)n(x,y) \rightarrow \ln(s(x,y)) = \ln(f(x,y)) + \ln(n(x,y))$. $\ln(n(x,y))$ is a high-frequency noise that can be filtered out using LPF. Signal is restored using exp(.) funcion. Draw a diagram:
   $\widehat{f}(x,y) = e^{LPF(\ln(s(x,y)))} = e^{LPF(\ln(f(x,y)))} e^{LPF(\ln(n(x,y)))} = e^{LPF(\ln(f(x,y)))} \approx f(x,y)$

2. $g(x,y) \rightarrow e^{g(x,y)}; e^{g(x,y)} = f(x,y) + n(x,y) \rightarrow \widehat{f}(x,y) = LPF(e^{g(x,y)})$. Then,
   $\widehat{f}(x,y) = LPF(f(x,y)) + LPF(n(x,y)) = LPF(f(x,y)) \approx f(x,y)$.

Question 2

1.

   a. 5 gray levels can be coded with $\lceil \log_2(5) \rceil = 3 \frac{bits}{pixel}$. Therefore, we need $3 * 10^4 = 30000$ bits to code this image

   b. Entropy of the source determines the lower bound on the average number of bits per symbol required to code the symbols for error free reconstruction.

   $$H = -\sum_{i=0}^{L-1} p(a_i) \log(p(a_i)) \frac{bits}{symbol}.$$

   $$H = -2 * 0.2 \log 0.2 - 2 * 0.1 \log 0.1 - 0.4 \log 0.4 \log 5 - 0.2 = 2.121928 \frac{bits}{symbol}$$

   Therefore we need $\lceil 2.121928 * 10^4 \rceil = 21220\ bits$.

   c.

level	1	2	3	4	5
a.	000	001	010	011	100
b.	10	0	1110	1111	110

2. $H = -\sum_{i=0}^{L-1} p(a_i)\log(p(a_i))\frac{bits}{symbol} = \frac{1}{2}\log 2 + 8*\frac{1}{16}\log 16 = \frac{1}{2} + 2 = 2.5\frac{bits}{symbol}$

$L_{avg} = \sum_{i=0}^{L-1} p(a_i)l(a_i)\frac{bits}{symbol} = \frac{1}{2}1 + 8*\frac{1}{16}4 = 2.5\frac{bits}{symbol}$

3. There is only one unique Huffmann code, where code length for every symbol is 2 (nothe that flipping bits and exchanging symbols does not change the code).

4. HINT1: $H = -\sum_{i=0}^{L-1} p(a_i)\log(p(a_i))\frac{bits}{symbol}$, $\sum_{i=0}^{L-1} p(a_i) = 1$. H is max for $p(a_i)$ if $\frac{\partial H}{\partial p(a_i)} = 0$ and $\frac{\partial^2 H}{\partial p(a_i)^2} < 0$. HINT2: try to prove that $H = -\sum_{i=0}^{L-1} p(a_i)\log(p(a_i))\frac{bits}{symbol} \le \log L$, using inequality $\ln(\frac{1}{x}) \ge 1 - x$, if $0 < x$.

Question 3

1.
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 5 & 5 & 5 & 5 & 6 \\ 3 & 3 & 3 & 5 & 5 & 5 & 5 & 6 \\ 3 & 3 & 3 & 5 & 5 & 5 & 5 & 6 \\ 3 & 3 & 3 & 5 & 5 & 5 & 5 & 6 \\ 5 & 5 & 5 & 5 & 5 & 6 & 6 & 6 \\ 7 & 7 & 7 & 7 & 7 & 7 & 6 & 6 \end{bmatrix}$$

2. Let n be the number of pixels in an image and n_k number of pixels in an image with intensity r_k.

a. FIRST PASS: $S_k = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n} = \frac{1}{n}\sum_{j=0}^{k} n_j \to S_l$. Since every pixel vith value r_k is mapped into S_l, we have that: $\sum_{j=0}^{k} n_j = \sum_{i=0}^{l} n_{S_i}$, $\forall l$ subjected to: $T(r_k) \to S_l$ and $T(r_{k+1}) \to S_{l+1}$.

b. SECOND PASS: $V_l = T(S_l) = \frac{1}{n}\sum_{i=0}^{l} n_{S_i} = \frac{1}{n}\sum_{j=0}^{k} n_j \xrightarrow{a.} S_l$. Therefore $T(S_l) \to S_l$, and a second pass of histogram equalization will produce exactly the same results as the first pass.

3. Frame rate is $30\frac{frames}{\sec}$ We are transmitting $3\frac{images}{frame}$: (1) right eye image, (2) left eye image, and (3) mounth image. There are 100 different images stored for each (1) right eye image, (2) left eye image, and (3) mounth image. For each ID number transmitted we need $\lceil \log_2(100) \rceil = 7\frac{bits}{image}$. Therefore, the bit rate is $30\frac{frames}{\sec} * 3\frac{images}{frame} * 7\frac{bits}{image} = 630\frac{bits}{\sec}$.

Question 4

1. Resulting bit allocation map:

2	1	0	0
3	2	0	0
4	3	2	1
5	4	3	2

2. In transform coding, an image is divided into blocks which are then transformed using DFT or DCT. The resulting coefficients inside the block are quantized independently (ecxcept for the JPEG DC component). Blocking artifact is a visual effect created due to the quantization of the DCT/DFT coefficients of the image blocks, where at lower bit rates the reconstructed image block boundaries become visible. In this context, DCT is better that DFT since DCT period is 2N (twice the block length), where the period of DFT is N (block length). In DCT based compressions, blocks appear less frequently and are less noticeable.

3. [*Note: you do not have to provide all the complexity details but important to point out that block computations are better/less expensive than whole image computations]*. The computational complexity of a transform coding is high. A DCT of a full $m \times m$ costs $O(m^3 \log^2 m)$. An DCT of $n \times n$ blocks costs $\frac{m^2}{n^2} O(n^3 \log^2 n) = O(m^2 n \log^2 n)$, and $O(m^2 n \log^2 n) < O(m^3 \log^2 m)$ for $n < m$. So it is cheaper to compute the transform for multiple smaller blocks that for the big image, the tradeoff being higher reconstruction error.

Question 5

1. In zonal coding, the coefficients with max variance over the image subblocks are retained. It results in the same zonal mask for all subimagrs. In threshold coding, location of the transform coefficients retained varies from one subimage to another. Only the largest transform coefficients for a subimage are kept

2. The quantization matrix defines the quantization step size. Step sizes are small for low frequancy components and large for high frequencies. The quantized divides the DCT coefficients by the corresponding quantums and then rounds to the nearest integer. In this process many high frequency components become zero and, therefore, easier to code.

3. Motion compensation is a video compression technique. The smallest difference between a current block in the current frame and a block from reference frame is DCT transformed and encoded using variable length codes. Motion vector represent a relative

displacement of the current frame w.r.t. reference frame. The computed motion vector is variable-length coded and transmited as an integral part of a data stream.

4. MPEG achieves its high compression rates by exploiting the fact that one video frame is very much like the frame that precedes it. I-frame (Independent frame) is compressed independently, and it is a reference point for motion estimation needed to generate susequent P- and B- frames. P-frame (Predictive frame) is the compressed difference between the current frame and the prediction of it based on the previopus I- or P- frame. B- feame (Bidirectional frame) is the compressed difference between the current frame and a prediction of it based on the previous I- or P- frame and next P- frame. These frames are significantly smaller that the full frames.

Question 6

State if the following statements are True or False.

a. Hauffman coding results in a lossy compression. FALSE

b. A more likely event has more self-information. FALSE

c. Transform coding is often used for loss-less image compression. FALSE

d. Transform coding reduces coding redundancy. FALSE

e. The DCT is the real part of the DFT. FALSE

f. Blocking artifact is less pronounced for the DCT than DFT in transform coding. TRUE

g. JPEG compression uses threshold coding. FALSE

h. JPEG compression standard uses DFT for compression. FALSE

i. Predictive coding reduces inter-pixel redundancies.TRUE

j. DCT closely approximates Karhunen-Loeve transform for a large class of images TRUE