ECE 634 Digital Video Systems – Spring 2015

Exam April 9, 2015

Name______________________________________

PUID ______________________________________

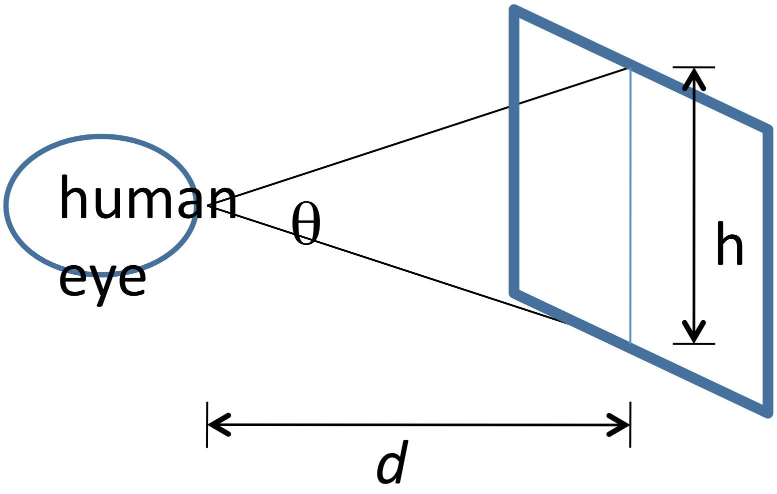
Instructions:

1. **Do not start until told to do so!**
2. **Write your name and student ID# above.**
3. **This is closed book exam, but you may have one sheet, ONE SIDE of notes**
4. **If extra paper is needed, use the back of the test pages**
5. **Cheating will not be tolerated. Cheating on this exam will result in, at minimum, a zero grade for the exam.**
6. **Continuing to write after the exam time has elapsed is considered to be cheating.**
7. **If you cannot solve a problem, be sure to look at the other ones, and come back to it if time permits.**
8. **Look over the entire exam before beginning! Note that problems toward the end of the exam have more points. Manage your time.**

Purdue University School of Electrical and Computer Engineering
ECE 634 Digital Video Systems, Spring 2015 (Prof. Amy Reibman)
Exam, 4/9/2015, 1:30-2:45pm
Closed book, 1 sheet of one-sided note allowed

---

1. (10 points). When designing a display system, to determine the necessary spatial resolution (number of lines M and number of pixels/line N), one should consider the display window size, the typical viewing distance, and the visual sensitivity to spatial frequency. Suppose the display size is *w*x*h* (in meters) and the viewing distance is *d* (in meters). We know that the human visual system is not very sensitive to spatial frequency above 10 cycles per degree. What is the minimum value of M and N you should use?



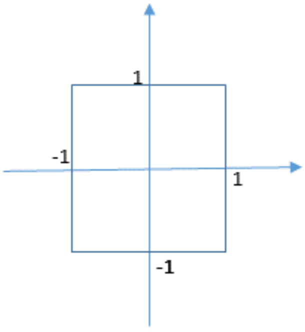
2. (15 points). Suppose a video is taken by a moving camera that is undergoing pan (rotation in the Y-axis) with rotation angle θ_y followed by zoom (change in focal length). The 3-D positions of any object point before and after this camera motion are related by

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}.$$

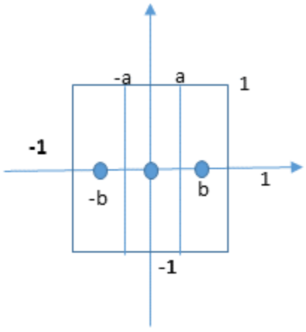
The camera focal length was changed from F to F’. Assume that the camera projection can be approximated by the perspective projection, i.e. $x=FX/Z$, $y=FY/Z$. Derive the 2D motion field induced by this camera motion.

3. (15 points). Consider a discrete memoryless source that has 4 symbols: a,b,c,d. The probabilities of each symbol are given as: p(a)=1/4, p(b)=p(c)=1/8 , p(d)=1/2 .
    a. What is the entropy of this source?
    b. Design a Huffman code for this source.
    c. Suppose your coder from part (b) outputs the bit string 0010001110101101011010110110. Decode the first 7 symbols using the coding scheme you designed in part (b).
    d. Compute the average number of bits per symbol that your encoder produces. Show your derivation.

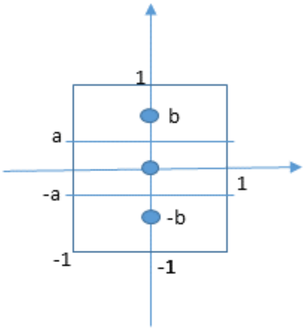
4. (20 points). Consider coding a 2-D random vector that is uniformly distributed over the region illustrated in Figure 4(a). Suppose you want to design a codebook with 3 codewords. One possible codebook construction (codeword locations and region partitions) is shown in Figure 4(b).
   a. Determine the values of *a* and *b* in Figure 4(b) that will minimize the mean squared error of the quantizer. Also, determine the corresponding minimal mean squared error.
   b. Another possible codebook configuration is shown in Figure 4(c). Is this codebook better, worse, or equivalent to that in Figure 4(b) if you choose the same value of *a* and *b*? Why?
   c. Sketch another codebook construction in Figure 4(d) with 3 codewords that may have lower mean squared error. You don't need to be precise, just provide a sketch and your reasoning why your codebook may be better.



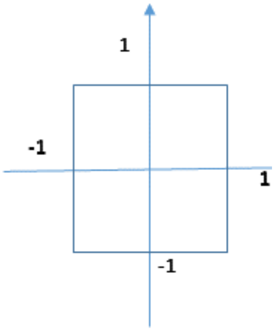
(a)



(b)



(c)



(d)

5. (20 points). Consider a predictive coding method in which pixels A,B, and C are co-located pixels in the frames at time instants t, t+1, and t+2, respectively. Suppose all pixels have variance σ^2 and the correlations between samples are $\rho_{AB} = \rho_{BC} = \rho$ and $\rho_{AC} = \rho^2$.
   a. Write the expression for the expected prediction error when B is predicted from A using a general predictor with weight α.
   b. Derive the optimal α to minimize the prediction error from (a). What is the coding gain relative to no prediction? What is the coding gain relative to prediction with α=1 (as is typically used in a video encoder)?
   c. Compare the following two scenarios to coding the sequence of pixels A, B, C. In the first scenario, C is predicted from A using unity prediction (α=1) and B is predicted using B=(A+C)/2. In the second scenario, B is predicted from A with unity prediction and C is predicted from B with unity prediction. What is the total prediction error in each scenario? For what values of ρ is the first scenario better than the second?

6. (20 points)

   (a) During this class, we often modeled a video as a stationary Gaussian process with variance σ^2 and correlation ρ between two adjacent samples. List 3 ways these models can be used to create effective designs for video compression, and list 3 ways video coding standards modify these designs to account for the fact that real video is not always well described by these models.

   (b) Describe what a video coding standard specifies and does not specify. Give 3 examples of parts of either encoder or decoder which are not standardized. Give 2 examples of parts of an ***encoder*** that ***are*** standardized.

   (c) Describe the evolution of motion representation and compensation from the video coding standards H.261, MPEG-1, MPEG-2, MPEG-4, and H.264. Be specific about the changes in how motion is described and what motion models are applicable in each of these standards.