Visual Perception: Midterm Exam
Total Points: 110
Due Date: 29 May, 2008

1. **Read the questions carefully. The points assigned to each question are indicative of the difficulty and length of their solutions.**
2. **Precise and to-the-point answers presenting simple and elegant solutions will be given more credit.**
3. **You are expected to work alone on the midterm and should discuss your answers with your classmates.**
4. **Please try to type in as much of the answers as you can. The equations, matrices and illustrations which are more time consuming to do in typing, it is okay to write them in hand.**
5. **Please hand me a hardcopy on the due date in class.**

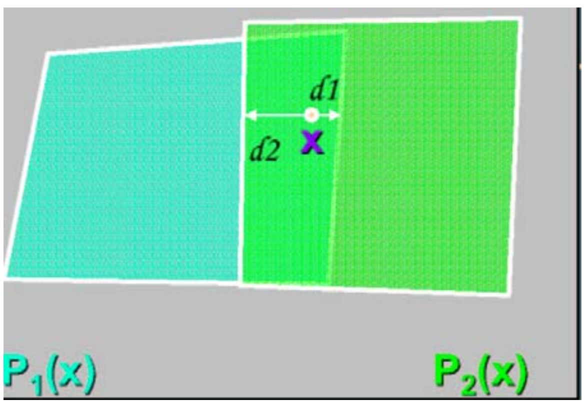
PLEDGE: I have not given/received any assistance to/from anybody in this exam.

Signature: ______________________________ Date: ______________

Name : ____

1. Explain why windows in a house, when viewed from the inside, appear transparent during the daylight hours, but act like mirrors at night. **[4]**

2. Explain how our perception of the Mach band would change if we had lateral excitation instead of lateral inhibition. **[4]**



3. The image on the left shows a display made of two projectors P_1 and P_2. At any pixel x, the contribution of the two projectors at this pixel are expressed by $P_1(x)$ and $P_2(x)$ respectively. To remove the high brightness in the overlap region, the intensity at any pixel x in this region is blended using the functions $A_1(x)$ and $A_2(x)$ from projector P_1 and P_2 respectively such that $A_1(x) + A_2(x) = 1$ and the combined intensity at pixel x from the two projectors is given by $A_1(x)P_1(x) + A_2(x)P_2(x)$.

Two types of function can be used.

(1) In the first, A_1 and A_2 are assigned as follows.
$A_1(x) = d1/(d1+d2)$
$A_2(x) = d2/(d1+d2)$

(2) In the second, A_1 and A_2 are assigned as follows
$A_1(x) = cos(d2/(d1+d2)*90)$
$A_2(x) = cos(d1/(d1+d2)*90)$

Which of these functions would look perceptually better? Why? **[4]**

4. We know that edge detection can be performed by the eye in different resolutions. Do you think that the illumination and the reflectance edges will show up in different resolutions? Justify your answer. **[3]**

5. You have your framed graduation certificate hanging on the wall. When you view if from a distance of 12 or more feet, you cannot see the letters clearly. As you approach the frame, slowly the letters become clearer and clearer. However, if you are within an inch or two, you find that the letters have again become blurred. Can you explain this with CSF? **[4]**

6. Consider two types of gamut mapping methods: a global method where all the colors are shifted to optimize the distance between them to be perceptually uniform and a local method which maps the out of gamut color only to assure that

they lie within the gamut. What are the pros and cons of both of these methods? What kind of applications do each of them cater to and why? **[5+5=10]**

7. Let a projector coordinate system be defined by *(s,t)* with origin *(0,0)* at the center of the projector. The projector shows a symmetric fall-off in the maximum luminance for each channel from the center to the fringes. The fall off is inversely proportional to the square of the distance of *(s,t)* from the center of the projector. (Consider the decrement to be proportional to square of the distance). The black offset does not vary spatially and is defined by the XYZ vector (X_B, Y_B, Z_B). Each channel has a quadratic input transfer function. If the color with maximum luminance for red, green and blue channel at the center of the projector is defined by the vectors (X_r, Y_r, Z_r) , (X_g, Y_g, Z_g) and (X_b, Y_b, Z_b), what would be the equation defining XYZ coordinates of the color for any input (i_r, i_g, i_b) at any pixel *(s,t)* of the projector. **[7]**

8. A display is specified by the following. The chromaticity coordinates of the three primaries: (0.6, 0.2), (0.2, 0.6) and (0.2, 0.1); chromaticity coordinates of the white point (0.3, 0.3); maximum brightness (X+Y+Z) of the display: 1000 lumens. Find the maximum luminance (Y) that can be displayed by each channel of this display. **[7]**

9. Give the following 3x4 camera matrix P and a 3D point in homogeneous coordinates **X** = [0 2 2 1], (a) fine the 3D Cartesian coordinate of X, (b) find the 2D Cartesian coordinate of the projection of X using the camera P. **[3+3=6]**

$$P = \begin{bmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{bmatrix}$$

10. An ideal pinhole camera has a focal length of 5mm. Each pixel is 0.02mm x 0.02mm and the principal point is at pixel (500,500). Pixel coordinates start at (0,0) at upper left corner of the image. What is the 3x3 intrinsic camera calibration matrix K for this camera? Assuming that the camera coordinate systems is coincident with the world coordinate system and the origin of the world coordinate coincides with the pinhole, what is the 3x4 extrinsic rigid body transformation between the camera coordinates and world coordinates? Combining the results from the above two questions find the 2D projection on the image plane of a 3D point (100, 150,800). **[8+2+4=14]**

11. The geometric property of a camera is defined by a 3x4 projection matrix P. Can this matrix P incorporate any lens distortion of the camera? Justify your answer. Can you recover the focal length of the camera from P? If so, give an algorithm to do so? **[3+6=9]**

12. You are given three cameras while estimating the 3D depth of a scene. Can you do better than having just two cameras? If not, justify your answer. If yes, state why and give an algorithm on how to achieve this betterment? **[5]**

13. The following camera projection matrix P gives the relationship between a 3D point (X,Y, Z) and its projector (u,v). How do you represent a scene point at infinity along the X-axis in the world coordinate using homogeneous coordinate? What is the image coordinates of the projection of such a point in both homogeneous and Cartesian coordinate? From the above, what can you say happens to vectors parallel to X axis in 3D world when projected on to camera images? **[2+ 4 + 3 = 9]**

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

14. You are asked to generate a panorama from a bunch of images automatically.

    a. If you want to solve this problem using homography, what should be your assumption? For what kinds of scenes will this assumption be true? **[5]**

    b. How would you use homography to solve the geometric calibration problem? Provide an algorithm. **[10]**

    c. What kind of artifacts can you expect at the common region between adjacent images and why? **[5]**