Visual Perception: Midterm Exam
Total Points: 75
Due Date: May 1, 11am

THIS IS A TAKE-HOME OPEN-BOOK OPEN-NOTES EXAM.
TIME= 2.5 HOURS.

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. **Please try to type in as much of the answers as you can. If there are equation and illustrations which are more time consuming to do in typing, it is okay to write them in hand.**
4. **Please hand me a hardcopy on the due date in class.**

You are bound by honor code to do the following.

1. **You are expected to work alone on the midterm and should not discuss your answers with your classmates.**
2. **You are supposed to spend no longer than 2.5 hours on the exam.**
3. **You are not supposed to take a peek at the questions ahead of the 2.5 hour you spend on the exam.**

PLEDGE: I pledge to have followed all the rules mentioned above.

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. **[3]**

   Weber law tells us that our ability to detect a difference is dependent on the strength of the stimulus. During night, the strength of the stimulus from inside the house is much more when compared to the strength of the stimulus outside and goes beyond the weber threshold of detection. Hence, all inside things can be seen in the reflections.

2. The color of the runway lights in airports is blue, and not red. Why? **[3]**

   Runway lights are usually used at night when our rods are in action due to scotopic conditions. Due to **the phenomenon of Purkinji shift** we are more sensitive to blue-greens at night. Hence, the runway lights are blue.

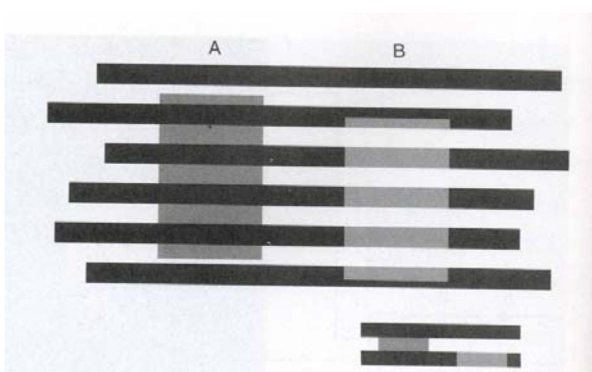
3. What will be the difference of our perception of Hermann's grid if we had lateral excitation instead of lateral inhibition (i.e a center surround receptive field that has inhibition in the center and excitation in the surround)? **[3]**

   With lateral excitation, the intersection of the corridors would appear lighter instead of darker. More excitation will come from whites than blacks creating this phenomenon.

4. Why do we see better in the dark if we avert/skew our gaze a little?**[3]**

   Our cones does not work well in scotopic conditions. There are no rods in the fovea where the image falls when we start directly at something. Averting our gaze makes the image fall on the periphery of the fovea which is rich in rods and hence we can see better.

5. In the image below, the gray rectangle on the right looks lighter than the rectangle on the left. But they are actually the same gray. Can you explain this by lateral inhibition? **[4]**

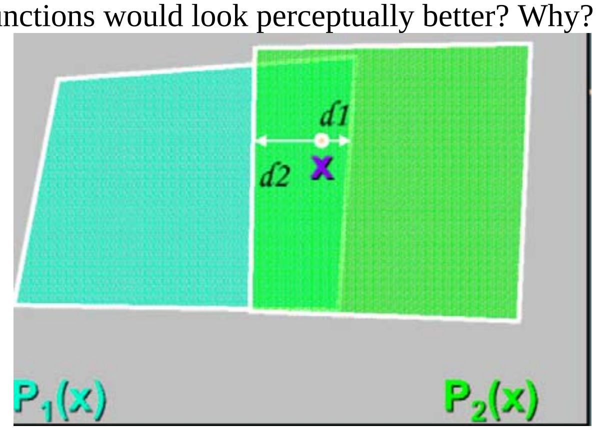


This is called the white's illusion and **cannot be explained by the lateral inhibition**. The lighter rectangle has lot of white surrounding it which should indicate higher lateral inhibition and hence a darker appearance. However, on the contrary, this rectangle appears lighter.

In fact, a theory of 'belongingness'explains this. By the structure, the lighter rectangle seems to belong to the foreground black horizontal bars while the darker rectangle seems to belong to the background. We perceive the contrasts of these rectangles with respect of where they belong and get this perception. **(The second part is for your information, I did not expect this in your answers).**

6. The image below 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.
   a. In the first, A_1 and A_2 are assigned as follows.
      $A_1(x) = d1/(d1+d2)$
      $A_2(x) = d2/(d1+d2)$
   b. In the second, A_1 and A_2 are assigned as follows
      $A_1(x) = cos(d1/(d1+d2)*90)$
      $A_2(x) = cos(d2/(d1+d2)*90)$

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



There was a slight mistake in the question. In (b), the correct functions should have been

$A_1(x) = cos(d2/(d1+d2)*90)$
$A_2(x) = cos(d1/(d1+d2)*90)$

If you compare these functions with that of (a), you will see that these have gradient continuity (C-1) as opposed to (a) which only has C-0 continuty and hence no gradient continuity. Hence, while **(a) will give mach band artifacts**, (b) will not. Hence (b) is better than (a).

Some of you caught this mistake and gave me correct answer. Some of you pointed out the mistake to me telling that even blending does not happen properly in (b)

since the ramp comes abruptly to 0 and hence will create mach bands. I have given full points to both.

7. 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. **[4]**

   We know that illumination edges are never sharp since there is no point light source in nature. They always come with a fuzzy umbra and penumbra region. Hence, **illumination edges are always blurrier than reflectance edges**. Coupled with this is the fact that illumination tend to change much slowly than reflectance. Hence, illumination edges will feature only in the lower resolutions.

8. 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]**

   Human CSF behaves like a **band pass filter**. It has a good sensitivity only when the frequency is between 3-30 cyles/degree. We the certificate is viewed from a large distance, the frequency subtended in the eye is greater than 30 cycles/degree and hence cannot be seen clearly. As we move closer to the frame, the frequency reduces to fall within the band of frequencies the human can see well. However, when the human is too close to the certificate , the frequency subtended in the eye becomes less than 3 cycles/degree and again clarity reduces.

9. 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 intensity of the display: 1000 lumens. Generate the matrix that converts the RGB coordinates for this device to the XYZ coordinates. What is the XYZ coordinates of the color generated by the RGB input (0.5, 0.75, 0.2) on this device? Find the maximum luminance (Y) that can be displayed by each channel of this display. **[6+2+2=10]**

   Let intensity of red, green and blue channel and the white formed by maximum intensity red, green and blue be Ir, Ig, Ib, and Iw respectively.
   From the information in the problem, Ir+Ig+Ib=Iw
   And Iw(0.3,0.3) = Ir(0.5,0.2) + Ig(0.2,0.6) + Ib(0.2, 0.1)
   The above gives you three equations which solved will give you Ir=250, Ig=350, and Ib=400.
   Since you know the chromaticity coordinates and I for each channel now, you can find X, Y and Z for each channel. For red – (150, 50, 50), for green – (70, 210, 70), and for blue – (80, 40, 280). So, the matrix M (in matlab format) is
   [150, 70, 80; 50, 210, 40; 50, 70, 280]

For the next question, you have to just post-multiply M with (0.5, 0.75, 0.2). The answer is (143.5, 190.5, 133.5). Finally, the luminance for red, green and blue channel you have already found as 50, 70 and 40 respectively.

10. 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. 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]**

The model would be
$[X, Y, Z](s,t) = i_r^2*[X_r\ \ Y_r/(s^2+t^2)\ Z_r] + i_g^2*[X_g\ \ Y_g/(s^2+t^2)\ Z_g] + i_b^2*[X_b\ Y_b/(s^2+t^2)\ Z_b] + [X_B\ Y_B\ Z_B]$
Some of you have given me the following answer for which I did not penalize you.
$[X, Y, Z](s,t) =(1/ s^2+t^2)[(i_r^2*[X_r\ \ Y_r\ \ Z_r] + i_g^2*[X_g\ \ Y_g\ \ Z_g] + i_b^2*[X_b\ Y_b\ \ Z_b] + [X_B\ Y_B\ Z_B])$

11. You are working in your office which receives good sunlight in the day time. Around dusk when it is getting dark, you find that it is difficult to read the text on the book. You light your table lamp and the text is again legible again. Explain this phenomenon with CSF. **[4]**

Human CSF changes with illumination. As the illumination decreases **the cut off frequency reduces.** Hence, frequencies which were seen clearly in high illumination will not be seen clearly when the illumination is dim. This is what happens as it gets dark. When the table lamp is put on, the illumination brightens increasing the cut off frequency where by the higher frequency text again becomes legible.

12. When we move our eye to look at a stationary scene, the objects do not seem to move around even though the image of the object on the retina moves. How can you explain this phenomenon? **[4]**

This can be explained by the **corollary discharge theory.** This theory states an exclusive OR connection exists in the higher visual processing areas that finds the exclusive OR of the movement of the image on the retina and the corollary signal due to the movement of the eye to signal movement of the object. When the eye moves to see stationary scene, the image of the retina and the eye both moves and hence the objects do not seem to move, but remain stationary.

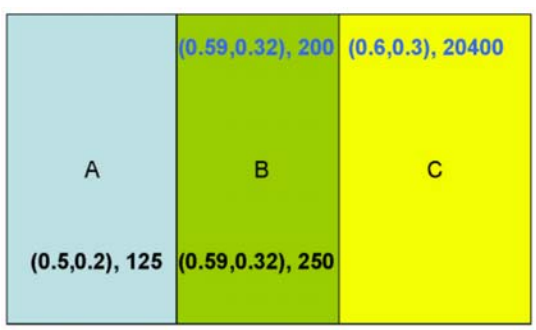
13. A person was presented an annular ring of luminance of 100cd/m^2 with a circular inset of luminance 50cd/m^2. In an adjacent region, he was presented with a circular inset of 150cd/m^2 and asked to adjust the luminance of the annular ring surrounding it so that it creates similar perception as that of the former pair. What luminance do you expect him to adjust to? **[3]**

    We know that we perceive the ratio of luminances at the edges. The ratio of the ring:inset = 2:1 in the first case. In the second case, we would try to match this ratio. For an inset of 150 cd/m2, this can be matched by a ring of 300 cd/m2.

14. The image below shows four different objects A, B, C and D, which have different reflectance. It also shows the gray scale luminance values at the edges. From this information how can you tell if the illumination is continuous or not? **[6]**

    If this image has only illumination changes, then the integration of the ratio of intensities across the reflectance edges would happen in a consistent manner even if we go via different paths. For example, to find C:A, if we go through D, we get the ratio of 9:2. Again, if we go through B also, we find the same ratio of 9:2. This means that the illumination is continuous.

15. The figure below shows an image segmented into three parts, A, B and C, based on the detection of edges between the regions. The chromaticity coordinates and the luminance in cd/m^2 on both sides of the edge between B and C are given in blue. The same for the edge between A and B is given in red. Based on this a person comments that there are more than one reason to believe that the former edge is due to illumination and the latter is due to reflectance. Do you agree with him? Justify your answer.**[4]**



Two observations support this claim. First, the chromaticity change between A and B is much more dramatic than between B and C telling that the former is a reflectance edge. The luminance change between B and C is very large which cannot happen in a reflectance edge, signifying the edge between B and C is probably an illumination edge.

16. You have difficulty in reading very small text on the screen. However, this difficulty reduces as you increase the size of the fonts. Explain this using CSF. **[4]**

    Human CSF is not sensitive at high frequencies (especially those greater than 30 cycles/degree). The small text in the screen falls beyond this frequency making them illegible. When the fonts are increased they come within the band of frequencies we are sensitive to making them legible.

17. Kay and McDaniel categorized colors in six categories corresponding to focal colors red, green, blue, yellow, black and white. How do you think would the sets corresponding to each of these categories look on the chromaticity diagram? Can you accommodate all of them on the chromaticity diagram? If not, why? Illustrate using diagrams.**[5]**

    All the color categories (note that I wanted you to mark the color categories roughly and not the focal colors) can be accommodated except black. Black and white fall in different places on the 3D XYZ space but they land in the same place on the chromaticity chart and hence they cannot be marked as two categories. Therefore, the different between them stem from the difference in their intensity and not in their chromaticity coordinates.