MIDTERM EXAM

Computational Perception (COMP 546)

Tues. March 12, 2019

Professor Michael Langer

STUDENT NAME: ____

ID: __________________

GRADE: ______/15

Instructions

- This is a closed book exam. No crib sheets.
- No electronic devices allowed.
- There are 15 questions, each 1 point, each covering one lecture.
- There is a mix of multiple choice questions and short answer questions.

  For the multiple choice, there are four choices on each question and exactly one is correct. If you circle the correct one, you get 1/1. If you do not circle any, you get 0/1. If you circle the wrong one, you will receive -0.2, that is, you will be penalized by 0.2 points. This penalty is meant to discourage you from guessing for questions in which you have no idea.

1. The diameter of a basketball hoop (rim) is about 50 cm. What is the *approximate* visual angle of the hoop in degrees, when viewed from a distance of 6 m ?

   Note that the hoop will appear as an ellipse. I am asking for the visual angle of the major axis.

   (a) 1 degree
   (b) 3 degrees
   (c) 5 degrees
   (d) 7 degrees

   SOLUTION:

   (c) $\frac{0.5}{6} * \frac{180}{\pi} = \frac{15}{\pi} \approx 5$

2. Recall the expression for the blurwidth in radians of a scene point at depth Z_0:

$$A(\mid \frac{1}{Z_o} - \frac{1}{Z_{focalplane}} \mid).$$

   Sketch the blurwidth of scene points as a function of scene depth *in diopters.*

   SOLUTION:

   The only thing you need to know in this question is what "diopters" means. You then need to plot the given function. See Exercises 2 Question 5b.

3. Suppose that a red object were placed against a grey background object and that this scene were viewed through the cyan filter similar to what is used in anaglyph glasses.

   Which of the two objects would appear brighter, and why? State your assumptions.

   SOLUTION:

   Assume the same simple model from the end of lecture 3. Let the light coming from the red object be represented by RGB values $(1, 0, 0)$ and let the light coming from the grey object be represented by $(.5, .5, .5)$, or any three roughly equal values that are less than 1 (i.e. grey, not white). Then the cyan filter multiplies channel-by-channel by (0,1,1) so the red object and grey object would have RGB = (0,0,0) and (0,.5,.5) respectively. *Thus the grey object would appear brighter.*

4. Why do retinal ganglion cells have both OFF-center/ON-surround and ON-center/OFF-surround receptive fields? i.e. Why not just one of these types ?

   SOLUTION:

   The cell response is a spike firing rate. Cells cannot have a negative firing rate. You can think of the OFF-center/ON-surround cell as encoding the negative responses of the ON-center/OFF-surround cell.

5. Which of the following is a *correct* statement about the receptive fields of cells in the *left half of area V1 in the brain*?

   (a) They are always tuned to some disparity, and sometimes the tuned disparity is 0.

   (b) They encode the image reaching the left half of the retina.

   (c) They define a retinotopic map with large receptive fields near the fovea and small receptive fields in the periphery.

   (d) They receive direct inputs from the retina.

   SOLUTION:

   (a) is incorrect, since not all cells are disparity tuned. **(b) is correct.** (c) is incorrect since cells in the fovea have small receptive fields and cells in the periphery have large receptive fields. (d) is incorrect since V1 cells receive inputs from the LGN, not directly from the retina.

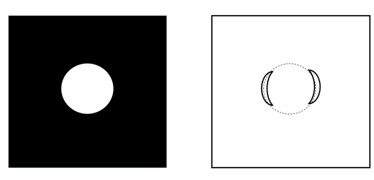
6. Consider a family of V1 complex cells which we can model using the outputs of Gabor cells of the form

$$\cos(\frac{2\pi}{N}k_1\ x)\ G(x,y,\sigma), \qquad \sin(\frac{2\pi}{N}k_1\ x)\ G(x,y,\sigma).$$

   Give a 2D sketch of the responses such complex cells to an image that consists of a white disk on a black background.

   Assume that the receptive field width of the Gabors is much smaller than the diameter of the disk.

   SOLUTION:



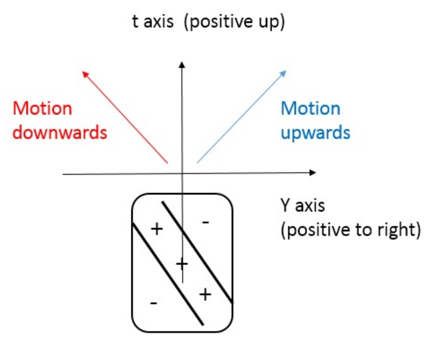
   The Gabors respond much more to vertical edges. So the cos and sin Gabors would have positive and negative responses near the crescent shaped region on the right, and zero everywhere else. However, the *complex cell* response would be only positive in the crescent shaped region on the right. **I also gave the point if you drew a rectangle or ellipse instead of crescent. However I only gave 0.5 points if you gave the responses of the cos and sin Gabors, which are simple cells, not complex cells.**

7. Sketch the YT receptive field of a cell that is tuned to downward vertical motion.

   Be sure to indicate the axes in your plot i.e. which Y and T directions are positive.

   SOLUTION:

   The receptive field should be for $t < 0$ only and the slope should indicate negative y velocity.



8. The motion constraint equation is

$$\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0.$$

   Give an example of intensity derivatives that would yield a motion constraint equation having the same "aperture problem" as the motion tuned cell in the previous question, namely a cell tuned to downward motion.

   Sketch this motion constraint equation in velocity space (v_x, v_y).

   SOLUTION:

   An example would be

$$\left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}\right) = (0, 1, 1).$$

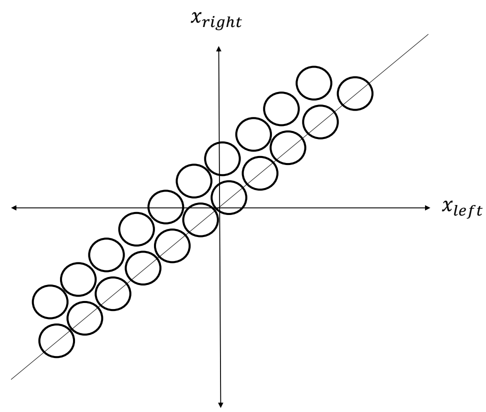
   which would give motion constraint equation $v_y = -1$. This is just a line parallel to the v_x axis in (v_x, v_y) space. **I gave the point if you drew such a line, even if you didn't give values for the derivatives.**

9. Many people who have normal vision in both eyes are nonetheless stereoblind, in that they cannot get any depth information from binocular disparities. Interestingly, some people are partly stereoblind, e.g. they can only perceive depth from crossed (+) disparities but not for uncrossed (-) disparities, or vice-versa.

   Using a disparity space image, sketch the receptive field distribution of cells of people who are stereoblind for crossed disparities only *i.e.* they cannot discriminate between depths of points with positive disparities $d = x_l - x_r > 0$.

   Hint: you will have to make an assumption about what causes the stereoblindness.

SOLUTION:



The point is that there are no cells tuned to positive disparities.

10. Which of the following statements about eye movements is *incorrect*?
    (a) The vestibulo-ocular reflex (VOR) operates independently of visual input.
    (b) VOR compensates for head rotation but not head translation.
    (c) Smooth pursuit eye movements are used to reduce the retinal motion of moving objects to zero.
    (d) Saccadic eye movements are used to direct the gaze to objects of interest, even if the objects are moving.

    SOLUTION:

    (b) is incorrect. VOR compensates both for translation and rotation. Many people answered (d), wish is arguably also incorrect if you consider that *after* a saccade is made to a moving object, the visual system needs to make smooth pursuit movement to keep the gaze on it. **I didn't take off the 0.2 point for (d).**

11. One method for discounting the illuminant in color constancy is to assume that the surface reflectances in the scene are grey on average. How could a vision system use this grey world assumption to infer surface reflectances?

    Note that reflectance is a fractional number (between 0 and 1).

    SOLUTION:

    The grey world assumption on reflectances would imply that the average color in the scene (defined say by the mean responses of each of the three cone types or mean RGB) would indicate the illuminant color (which doesn't vary with x, y). e.g. for the red channel:

$$mean_{x,y} I_R(x,y) = Illuminant_R \ * \ mean_{x,y} Reflectance_R(x,y)$$

    The visual system could estimate reflectance of each point by normalizing by the mean of each channel:

$$\left(\frac{I_R(x,y)}{mean_{x,y} I_R(x,y)}, \frac{I_G(x,y)}{mean_{x,y} I_G(x,y)}, \frac{I_B(x,y)}{mean_{x,y} I_B(x,y)}\right)$$

    which would give

$$\left(\frac{Reflectance_R(x,y)}{mean_{x,y} Reflectance_R(x,y)}, \frac{Reflectance_G(x,y)}{mean_{x,y} Reflectance_G(x,y)}, \frac{Reflectance_B(x,y)}{mean_{x,y} Reflectance_B(x,y)}\right).$$

    But the grey world assumption means that the denominators are all equal, so the triplets are just scaled reflectances (and note the average of each of these triplets is 1). Multipying by another scale factor would be needed to ensure all values are less than 1.

    I gave 0.5 for the first equation above, and another 0.5 for mentioning a normalization needed to bring values to 0 to 1 range. I didon't require the full explanation as above.

12. What is the mathematical model typically used for describe shading on a sunny day? Sketch an example of a shaded cylinder and indicate the direction of the sun.

<u>SOLUTION:</u>

$$I(x, y) = \mathbf{N}(x, y) \cdot \mathbf{L}$$

I did not expect a beautiful sketch. I just wanted a cylinder which has constant intensity parallel to the axis (where the normal is constant). The light source would be on the brighter side. See slide 3 from lecture 12. .

13. Contrast detection thresholds for 2D sinusoidal intensity images which are shown in the fovea are smallest for frequencies around 4 cycles per degree. Which of the following is *correct* about contrast detection thresholds in *peripheral vision*?

"The minimum occurs at a _____ frequency, and this minimum contrast detection threshold in the periphery is _____ than in the fovea."

(a) higher, larger
(b) lower, larger **SOLUTION**
(c) higher, smaller
(d) lower, smaller

14. The motion constraint equation is based on an assumption that the image intensities are translating with velocity (v_x, v_y). Assuming each image frame has additive Gaussian noise σ_n^2, which of the following would be a valid model of a likelihood function for estimating (v_x, v_y) ?

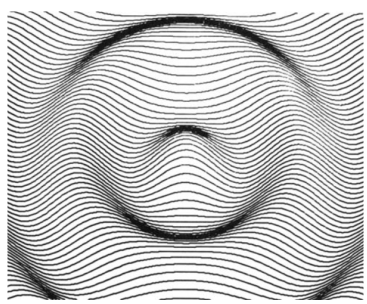
(a) $\Pi_{x,y}\ e^{-\ |\ (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}\ )\ |^2/2\sigma_n^2}$

(b) $e^{-\Pi_{x,y}(\frac{\partial I}{\partial x}v_x + \frac{\partial I}{\partial y}v_y + \frac{\partial I}{\partial t})^2/2\sigma_n^2}$

(c) $e^{-|(v_x,v_y)|^2/2\sigma_n^2}$

(d) $e^{-\sum_{x,y}(\frac{\partial I}{\partial x}v_x + \frac{\partial I}{\partial y}v_y + \frac{\partial I}{\partial t})^2/2\sigma_n^2}$ **SOLUTION**

15. In the image below, one tends to perceive the central region as a local hill rather than a local valley, whereas if one turns the page upside down the opposite tends to occur. How should one explain this tendency using Bayesian theory?



(a) there is a larger likelihood for hills than for valleys;
(b) there is a larger prior for hills than for valleys;
(c) there is a larger likelihood for upward slopes (floors) than for downward slopes (ceilings);
(d) there is a larger prior for upward slopes (floors) than for downward slopes (ceilings); **SOLUTION is (d)**