Questions

1. One interesting observation that is commonly made in the winter and on overcast (dense cloud) days is that the snow appears to be brighter than the sky. Specifically, when you stand in a snow covered field and you compare the ground below the horizon to the grey sky above the horizon, the ground seems to be brighter.

Why does the snow seem brighter than the sky? Doesn’t this defy the laws of physics?

2. Consider a surface whose depth varies sinusoidally in the $x$ direction but is constant in the $y$ direction. Suppose the surface is defined by random dots, displayed on a stereo monitor, such that corresponding dots in the left and right eye images produce retinal disparities

$$d(x, y) = a \sin(kx).$$

Describe a psychophysical experiment that measures

(a) disparity sensitivity thresholds as a function of $a$, for a fixed value of $k$.

(b) disparity sensitivity thresholds as a function of $k$, for a fixed value of $a$.

(c) The thresholds you compute in both (a) and (b) depend on both $a$ and $k$. Does this mean that the threshold values will be the same in the two cases?

3. Consider the claim that, “for surfaces that are near to the eye (say 30 cm), thresholds for slant angle from stereo increase as surface slant increases”.

(a) Explain what this claim means by describing a particular psychophysical task and a subject’s performance in this task.

(b) Argue how this claim could be a result of Panum’s fusional area.

4. (a) Suppose we have a likelihood function $p(I|S)$ and a prior $p(S)$, and that both are Gaussians. How can we combine them to estimate the maximum posterior for $S$?

(b) What if you have a prior and two likelihood functions, for example, you have texture and stereo cues that give you (conditionally) independent likelihoods and you also have a prior? How do you combine the weights so that you can compute a maximum posterior probability of $S$? You may assume:

$$p(S|I_1, I_2) = \frac{p(I_1|S)p(I_2|S)p(S)}{p(I_1, I_2)}.$$

5. One might expect the prior $p(x, y, d)$ on binocular disparities to have its peak at $d = 0$ for any $(x, y)$ near the center of the fovea(s), i.e. near $(x, y) = (0, 0)$, since the eyes tend to be aligned so that the center of fovea has zero disparity.

Would you expect the prior to also have its peak at $d = 0$ for other $(x, y)$ values? If not, how might the prior depend on $(x, y)$?
6. In lecture 17 p 2, when we discussed the noisy version of the motion constraint equation, I claimed that for any data values $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$, the expression
\[
\sum_{(x,y)} \left( \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} \right)^2
\]
is a quadratic function of $v_x, v_y$ and hence has a single minimum. Justify this claim.

Hint: this question has nothing to do probabilities or noise. Rather, it is just to make sure that you understand the claim.

Solutions

1. Although the sky might appear to be uniformly bright, in fact it is not. The zenith (high overhead) directions of the sky are about 3 times brighter than the horizon on an overcast day. Given this fact, consider how the luminance from the sky compares to the luminance of the light reflected from the snow. Snow is a diffuse reflector (i.e. on a sunny day, you don’t get highlights on the snow). So the luminance of the snow is a weighted average of the luminances from the different directions of the sky. So, the snow’s luminance is between the sky’s zenith luminance (max) and its horizon luminance (min) i.e. the snow has a greater luminance than the sky’s horizon.

The transition from the (darker) horizon to (brighter) zenith happens over 90 degrees of visual angle, which is a very low spatial frequency. Our visual systems have very poor contrast sensitivity at low spatial frequencies, so we don’t see this transition. i.e. The sky looks uniformly bright, from horizon to zenith i.e. the sky looks uniformly darker than the snow. That is the illusion!

2. (a) For any fixed amplitude $a$, the task could be to compare a standard surface
\[
d(x, y) = a \sin(kx)
\]
to a test surface
\[
d(x, y) = (a + \Delta a) \sin(kx)
\]
and decide which surface has greater amplitude in the depth variations.

There would be a psychometric function for each value of $a$. It would show the percentage of responses “test has greater amplitude than standard” as a function of $\Delta a$. The threshold value of $\Delta a$ would be defined as the value of $\Delta a$ where the response occurs some percentage (say 75) of the time.

Varying $a$ would give a set of thresholds.

(b) Consider a standard curve
\[
d(x, y) = a \sin(kx)
\]
and a test curve
\[
d(x, y) = a \sin((k + \Delta k)x)
\]
and the task of deciding which has greater frequency.
For a fixed $a$ and $k$, the psychometric function would plot the percentage response “test has greater frequency than standard” as a function of $\Delta k$. Define the threshold value of $\Delta k$ to be the value where the response was given some percentage (say 75) of the time. Varying $k$ would give a set of thresholds (for the fixed $a$).

(c) The thresholds you compute in both (a) and (b) depend on both $a$ and $k$. However, these thresholds in (a) and (b) are different since the task in (a) is to discriminate amplitudes and the task in (b) is to discriminate frequencies.

3. (a) The task is to distinguish the slopes (measured in slant angle) of different planar surfaces. Slant angle can go from $\theta = 0$ (frontoparallel) to 90 degrees (infinite slope). For each standard slant $\theta$, the task is to decide whether this standard is more or less slanted than a test slant $\theta + \Delta \theta$, where $\Delta \theta > 0$. Alternatively, both slants could be shown (say one after the other, in random order) and the subject would be asked to identify which was more slanted.

In the latter case, the psychometric curve could plot percent correct as a function of $\Delta \theta$. Threshold could be the $\Delta \theta$ such that subject was correct 75 % of the time.

(b) If the surface is more slanted, then the range of disparities covered by any visual angle will increase. But for nearby viewing, the range of disparities would be very large when the slant is large. In this case, only the points at the same distance as the vergence point will have disparities small enough to be fused i.e. Panum’s fusional area. But if fewer points are fusible, then there is less stereo information available for the visual system to estimate the slant.

4. (a) Assume the prior and likelihood have Gaussian shapes (ignore the constants) with means $S_1, S_2$ and standard deviations $\sigma_1$ and $\sigma_2$, respectively.

Then, similar to what was done in lecture 17, we can write the product of the likelihood and prior as

$$p(I|S)p(S) = e^{-\frac{(S-S_1)^2}{2\sigma_1^2}} e^{-\frac{(S-S_2)^2}{2\sigma_2^2}}$$

(times a constant we don’t care about). This has the same form as what we saw in lecture 17 where we had the product of two likelihoods, namely the product of two Gaussians.

To maximize the above product, we want to minimize the sum

$$\frac{(S-S_1)^2}{2\sigma_1^2} + \frac{(S-S_2)^2}{2\sigma_2^2}$$

Now take the derivative of $p(I|S)p(S)$ with respect to $S$, and set the derivative to 0. This gives $S = w_1 S_1 + w_2 S_2$ where the weights $w_1$ and $w_2$ are the same expressions given in lecture 17.

(b) Letting $S_p$ and $\sigma_p$ be the Gaussian parameters for the prior. We want to maximize

$$p(I_1|S)p(I_2|S)p(S) = e^{-\frac{(S-S_1)^2}{2\sigma_1^2}} e^{-\frac{(S-S_2)^2}{2\sigma_2^2}} e^{-\frac{(S-S_p)^2}{2\sigma_p^2}}$$

so we want to minimize

$$\frac{(S-S_1)^2}{2\sigma_1^2} + \frac{(S-S_2)^2}{2\sigma_2^2} + \frac{(S-S_p)^2}{2\sigma_p^2}.$$
Again take the derivative with respect to $S$ and set it to 0. This gives

$$\frac{S_1 - S}{\sigma_1^2} + \frac{S_2 - S}{\sigma_2^2} + \frac{S_p - S}{\sigma_p^2} = 0$$

and so

$$S = \left(\frac{S_1}{\sigma_1^2} + \frac{S_2}{\sigma_2^2} + \frac{S_p}{\sigma_p^2}\right) \cdot \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \frac{1}{\sigma_p^2}\right)^{-1}$$

This is the estimate of the maximum of the posterior. Note that this is of the form

$$S = \sum_{i=1}^{3} w_i S_i$$

where

$$w_i = \frac{\sigma_i^{-2}}{\sum_i \sigma_i^{-2}}.$$

5. In an outdoor scene, the region above the fovea ($y > 0$) tends to contain the sky and in that case its disparity be negative (and $l$ equal to the vergence angle). The region below the fovea ($y < 0$) will tend to be on the ground and closer than the object seen at the fovea, and hence its disparity will tend to be positive.

Another way to reach this conclusion is to note that there is prior for viewpoint from above (or “floor” slant instead of “ceiling”) which leads to the same prior on depths (or disparities) above or below the fovea.

6. The summation is of the form

$$\sum_i (a_i v_x + b_i v_y + c_i)^2$$

If we expand each of the terms in the sum, and sum them up, we get an expression of the form

$$\alpha_0 v_x^2 + \alpha_1 v_y^2 + \alpha_2 v_x v_y + \alpha_3$$

which is a second order polynomial in $v_x, v_y$, i.e. quadratic. Moreover, inspecting the summation, we see that $\alpha_0 > 0$ and $\alpha_1 > 0$. Hence the sum goes to $+\infty$ when either $v_x \to \pm\infty$ or $v_y \to \pm\infty$ because each of the terms in the sum are squared (hence positive). Thus it has a minimum for finite values of $(v_x, v_y)$. But since it is quadratic, the minimum is unique.