# Questions

1. (a) Consider two Gaussian functions $G(x, \sigma_1)$ and $G(x, \sigma_2)$. Assume $\sigma_1 < \sigma_2$. Is the following DOG ON-center/OFF-surround or OFF-center/ON-surround ?

$$DOG(x, \sigma_1, \sigma_2) = G(x, \sigma_1) - G(x, \sigma_2)$$

   (b) Sketch the cross-correlation of ON-center OFF-surround DOG function with an image of a bright disk on a black surround. Assume the bright disk is much smaller than the DOG. The DOG and disk are circularly symmetric, so you can sketch your response in 1D.

   (c) Same as (b), but now suppose the bright disk is much larger than the DOG.

   (d) Suppose you have an image $I(x, y)$ which is a vertical white line at $x = x_0$ and otherwise is black. Sketch the result of the 2D cross-correlation $DOG(x, y, \sigma_1, \sigma_2) \otimes I(x, y)$. Be sure to label where $x = x_0$. Also, you may assume that the line has some finite thickness to it. (Otherwise, the math would say there is no response.)

   (e) How would your answer change if the image were a dark line on a white background?

2. The *blind spot* is a particular location in each retina where the axons from all the ganglion cells are gathered and bundled together, where they then head off toward the brain (as the optic nerve).

   To find the blind spot, close your left eye and look at the left dot below, using your right eye. (Or close your right eye, and look at the right dot below using only your left eye.) Move your head toward or away from the page until the other dot disappears.

   ■ ■

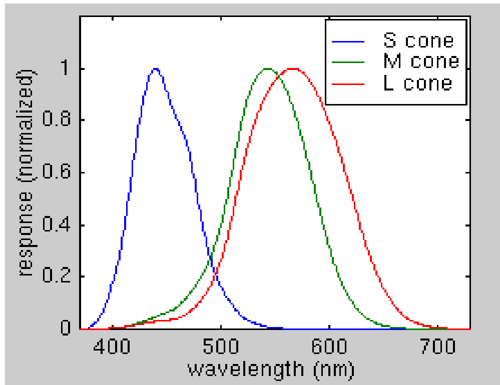
   How far away is this spot in the periphery (in deg of visual angle)

   If you can't find it, let me know. Everyone eye has one.

3. In the lecture I showed a plot of the density of photoreceptor cells in the retina and it was in units of cells per $mm^2$. For example, near the center of the fovea the density peaked at about 140,000 per $mm^2$.

   What is the visual angle that corresponds to 1 $mm^2$ ? What is the sampling rate (pixels/degree) that corresponds to the maximum shown in the plot, namely 160,000 pixels/$mm^2$.

4. (a) For color blind people who are missing their L cones, red colors tend to look dim. An example is red traffic lights. Explain why this dimness problem occurs.




Note: a person missing L cones doesn't necessarily have fewer photoreceptors. For simplicity, assume the L cones are replaced by M cones.

(b) What (if any) darkening effect would you expect for color blind people who are missing their M cones?

(c) A person who is missing the S cones is likely to confuse light blue (say RGB = (.5, .5, 1)) with which of the following colors? You may assume that there are plenty of other colors in the pattern with which to make a comparison.

For any that you select, briefly explain why.

- white
- grey
- black
- yellow
- red
- green

You may assume that we are considering a situation in which the light blue color patch is seen among many other color patches.

# Solutions

1. (a) ON-center, OFF-surround

   (b) If the disk is very small, then the image can be approximated by a single bright pixel. So as you move the DOG template across the single bright pixel (that's what cross-correlation does), the output will just be the value of the DOG at the pixel. So, the response will just correspond to the weights of the DOG at different positions of the DOG: the output will look like the DOG itself. (We will see this result again later in the course when we examine linear systems and in particular delta functions.)

   (c) The response will look like this, as the DOG is moved to different positions across the disk:

   ```
   00000000000---+++0000000000000+++---00000000
   ```

   where 0 is zero, - is negative, and + is positive outputs.

   What's going on here is that when the DOG first overlaps the disk, it will be the OFF region of the DOG that will overlap. So the output will first be negative. Then as the DOG goes further inwards into the disk, eventually the ON regions of the DOG will lie in the bright disk but some of the OFF region of the DOG will not. So the overall output will be positive. Since the bright disk in this question is assumed to be much bigger than the DOG, there will be positions of the DOG where it lies entirely within the disk and the output will be 0. As the DOG moves further and starts partly overlapping the opposite edge of the disk, the opposite pattern of + and - will occur.

   (d)
   ```
                         x0 is here
                            |
   000000000000000000---+++---00000000000000000000000
   000000000000000000---+++---00000000000000000000000
   000000000000000000---+++---00000000000000000000000
   000000000000000000---+++---00000000000000000000000
   000000000000000000---+++---00000000000000000000000
   000000000000000000---+++---00000000000000000000000
   000000000000000000---+++---00000000000000000000000
   ETC
   ```

   Notes:

   - If the line is long and extends for the whole image, then the image is constant in the y direction. So the result of the convolution would be constant in the y direction too. (I have only shown a limited y range above)
   - You may be wondering why the central region of `+++` is exactly three 'pixels' wide. The three pixel width is not meant to be taking literally. I just meant that there is a region where the linear response is positive.
   - You might think that the central region should be `+0+`, that is, there should be a 0 response in the middle where the line cuts across the center of the receptive field. To understand why the linear response is positive at that location, think about having a slightly thickened line which just covers the central region of the DOG, i.e. up to the

point where the DOG transitions from `+` to `-`. Such a line would produce a positive response from the DOG since the positive part of the DOG would have a greater total weight than the negative part of the DOG (within the region containing the thickened line). As you shrink the line down to 0, the response of the DOG would indeed shrink to 0 though.

(e) It would flip sign:

```
000000000000000000+++---+++0000000000000000000000
000000000000000000+++---+++0000000000000000000000
000000000000000000+++---+++0000000000000000000000
000000000000000000+++---+++0000000000000000000000
etc
```

2. The blind spot is roughly 15-20 deg away from the viewing direction. That is, the distance between your eye and the screen (or page) should be roughly 3–4 times the distance between the squares.

3. Assuming the eyeball is roughly 2 cm long. Then 1 mm on the retina would define a visual angle of $\frac{1}{20}$ radians, or about $3 \times 3$ degrees. Assuming for simplicity a square pixel grid, we have 160,000 pixels is a grid of $375 \times 375$ or about 125 per degree of visual angle in each row or column of 'pixels'.

4. (a) A red traffic light has most of its energy in the longer wavelengths, say beyond 600 n. The color blind person that is missing L cones would not be sensitive to such long wavelengths and so their M cones would have relatively little response. So the red light would appear dimmer than a green or blue light that had the same physical intensity.

   (b) You might try to apply the same argument and say that green traffic light (middle frequencies dominating) would look dim. However, note that the M cone receptor's sensitive region is almost entirely contained within the L cones sensitive region. So spectra that give a good response to the M cones would also give a good response to the L cones. (Of course, in either case, the person would have trouble distinguishing red and green as colors. The questions here are about brightness, not hue.)

   (c) Light blue would be confused with grey or yellow. Why? Missing the S cones means that colors cannot be compared on the yellow-blue opponency dimension. So, yellows would just appear neutral – they would stimulate both L and M. 'Black' and 'white' are incorrect since the R and G channels are .5 and so the light blue would appear darker than white and brighter than black. 'Red' (and 'green') are incorrect since these colors generally require an larger R than G ( or larger G than R).