COMP 546 Assignment 2 Winter 2018

Prepared by Prof. Michael Langer

Posted: Wednesday, Feb 14, 2018
Due: Thursday, March 1, 2018 <u>at 23:59</u>.

The same general instructions apply as in Assignment 1, except for the late policy which is replaced as follows:

> Late assignments will be accepted up to only 3 days late. They will receive a late penalty of 20 percent of your grade *per day*. If your assignment is submitted anytime on Friday (from 1 minute up to 24 hours late), then this is considered 1 day late and you would be penalized 20 percent of your grade. For example, if your raw grade were 70/100 then you would receive a 14 point penalty and so your grade would be reduced to 56/100.

Let's get started!

This assignment will give you some hands on experience with modelling simple and complex cells using Gabor functions. We examine various problems of binocular disparity estimation and monocular motion estimation.

1. (20 points)

   The Matlab script **disparity.m** defines a random dot stereogram with a central square in front of a background plane. The central region has some disparity in pixels. The surround has disparity 0. The script creates an analglyph figure for visualization purposes. You can borrow red-cyan glasses from me (ML) so you can experience the depth effect.

   Add code to the script to filter the left and right images with families of vertically oriented Gabor cells (the cosine and sine Gabors are models of simple cells) which are combined to give binocular complex cells. Each binocular complex cell has a peak sensitivity at some disparity. A range of disparities is given in the starter code. The binocular complex cells should be defined by the shifting model that I presented on pages 3-4 of the lecture 6 notes.

   When you've computed these responses, Figure 2 subplots will show responses of the binocular cells for a various disparity tunings.

   Briefly discuss the variations in responses *within and between* subplots. In particular:

   - Why are the responses not uniform in the center square (or background) region of each subplot?
   - For the center squares, how do the mean responses vary across different disparity-tuned cells? How (if at all) do these variations across disparity relate to the underlying wavelength of the Gabor?
   - Same as previous question, but for the surround square.

   To support your discussion for the latter two questions, *submit the figure* (generated by the starter code) that shows two 1D plots of the mean responses of each family of disparity tuned cells as function of the tuned disparity. One of the 1D plots shows the mean response of each family over the pixels in the center square region and the other plot shows the mean response of each family over the surround region. If you have constructed your Gabors correctly and if you are computing mean responses over the correct regions, then you should find that different disparity tuned cells have different mean responses for the center square versus surround.

2. (10 points)

   Using the same 1D plotting code as in the previous question, now consider (a) the responses of a family of diagonally (45 degrees) oriented Gabors and (b) a family of horizontally oriented Gabors. The new Gabors should have the same wavelength as the original, but should be rotated by 45 and by 90 degrees from the original.

   Do the responses of the vertically oriented Gabors from Q1 provide more precise information about (horizontal) disparity than the diagonally or horizontally oriented Gabors? Briefly discuss.

3. (20 points)

   a) Consider a different approach that measures the squared *differences* of Gabor responses rather than the squared *sums* of Gabor responses (see lecture 6 page 5). Recall the idea that if an image region in one eye is a shifted version of the corresponding region in the other eye, then a Gabor at some position in one eye should have the same response as the suitably shifted Gabor in the other eye, and so the difference in responses of the two Gabors should be zero. *Compute the squared differences as outlined in the lecture and give 1D 'mean response' plots similar to Q1.* Verify that the minima in responses occurs at the correct disparity, as claimed in the lecture.

   b) Read the following paper which is included in the zip directory:

      *"*Responses of primary visual cortical neurons to binocular disparity without depth perception*"* (B. Cumming and A. Parker, Nature 1997)

      Manipulate your random dot stimuli in a similar way to what they did in the paper, namely use anticorrelated image pairs rather than correlated.

      Compare the 1D mean responses of the correlated versus anticorrelated pairs. Are the results analogous to what they show in their Figure 2 ?

      Briefly discuss any relationships you notice between the plots here versus in part (a).

4. (10 points)

   Consider a slightly different geometry in which the "central square" now is an empty hole (or window) through which a background surface is partly visible. Assume that the background 'surface' that is seen through the hole is as bigger than the hole, so only part of the background is visible through the hole.

   We are not asking to make a stereo pair for this new configuration. Instead, just *draw the disparity space image* (lectures 6 and 10). It is enough to draw a 1D slice that cuts across the hole. Assume in your drawing that the foreground surface containing the hole has disparity = 4 and background seen through the hole has disparity = 0. That is, the eyes are verging at a point on the background, which is seen through the hole. Indicate which regions of the foreground and background are visible to both eyes, or only to the left or right eye, or to neither eye. You may draw the sketch by hand, take a JPG photograph of it, and insert it into your PDF.

   Hint: The background will contain binocular and monocular points, as well as points that are visible to neither eye. These should be labelled.

5. (20 points)

This question and the next considers *monocular motion* sensitive complex cells. For simplicity, let the cells temporal receptive field be just two frames (time t-1 and t). This allows us to implement the cell responses using similar code to the binocular case, where `left` and `right` for binocular cells become t-1 and t for motion. One difference for the motion case is that we assume the displacement over time of the receptive field is normal to the preferred orientation the cell. Note that the displacement $(\Delta x, \Delta y)$ in one frame can also be written (vx, vy) pixels per frame, and this question will use displacement and velocity interchangeably.

Consider four families of motion cells.

- The first is defined by vertically oriented sine and cosine Gabors whose receptive field at time t is displaced from its receptive field at time t-1 by $(\Delta x, \Delta y)$ = (4,0) pixels. This family of cells is exactly analogous to binocular cells tuned to horizontal (x) disparity of 4 pixels.
- The second and third family of cells have preferred spatial orientations in directions (1,1) and (1,-1), respectively. In particular they are defined by displacements (2,-2) and (2,2), respectively, which we can equivalently call their normal velocity. Verify for yourself that the motion constraint lines defined by their normal velocities pass through $(vx, vy) =$ (4,0).
- The fourth family is defined by horizontally oriented Gabors whose receptive fields at time t-1 are identical to their receptive field at time t, so that the cells prefers no motion in the direction normal to their preferred (horizontal) orientation. Verify for yourself that the motion constraint line defined by their normal velocity is $vy = 0$, which also contains velocity $(vx, vy) =$ (4,0).

Make a 2D plot of the mean responses of each of these four families of motion tuned cells to a pair of random dot images (similar to previous questions). The plots should show responses to a range of image velocities, i.e. displacements between the two frames:

$$(vx, vy) \in \{-10, \dots, 10\} \times \{-10, \dots, 10\}.$$

Use `imagesc` for the plots and submit four plots, namely one for each of the four family of cells. To generate the two-frame motion, use the `circshift` operator. Note that this operator allows you to shift by integer steps only, which is fine, given the values in this question.

6. (20 points)

In the lecture, you learned about motion constraint lines and the intersection of constraints solution to the aperture problem. However, those arguments were for an oversimplified situation in which one measures a local intensity gradient. In the previous question we did not measure local intensity gradients, but rather we measured Gabor filter outputs. Here you will consider the relationship between the intersection of constraints arguments given in the lecture and the Gabor filter responses.

Make a 5x4 table of the mean responses of the four families of cells in the previous question to several actual velocities (vx, vy) in pixels per frame, namely (4,0), (-4,0), (0,4), (0,-4), (0,0).

In the lecture I briefly described how MT cells can be defined by taking sums of V1 motion sensitive cells. Using the data from your plots and/or the table, briefly describe how the responses of the motion sensitive complex cells could be used to estimate the true velocity. The argument you should give here is similar to what I gave in the lecture, but here you should ground this argument on your plots and data.

Technical heads up: Matlab uses [row,column] indexing which is standard for matrices – see any linear algebra book. So the displacement due to velocity (vx, vy) will be indexed by [vy_index, vx_index]. Also note that the row index increases downward on the page, so the top left row of a matrix has row index 0.