Questions

1. Suppose the visual system estimates the direction of heading by estimating the velocity vectors at various points throughout the visual field and then finding a direction in the visual field from which all velocity vectors point away. In particular, assume there is no rotation component of the motion field.

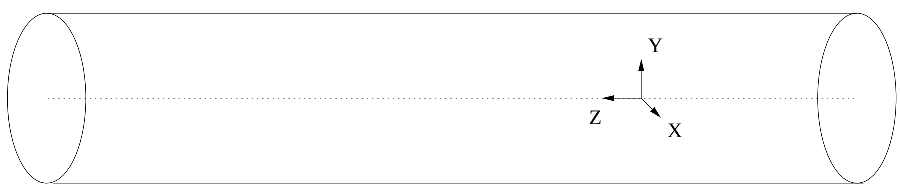
   Is the estimated motion field and direction of heading enough information to estimate the depth map ? Hint: see lecture 9 equation (5) on page 4.

2. Suppose an observer is moving directly forward on a ground plane and the observer's eye is a height h above the ground. What is the image velocity field (v_x, v_y) ?

   Hint: I give the answer in the lecture notes. To do this question, you need to identify two equations and substitute one into the other. Which equations?

   Assume for simplicity that the scene consists only of the ground i.e. no objects.

3. Suppose an observer is moving forward through a long hollow cylindrical tunnel of radius h, such that the optical axis coincides with the axis of the cylinder, and the projection plane is $Z = f$:

   (a) What is the motion field as a function of (x, y)?

   (b) Sketch this motion field.

   Hint: To do this question, you need to first do the previous one.

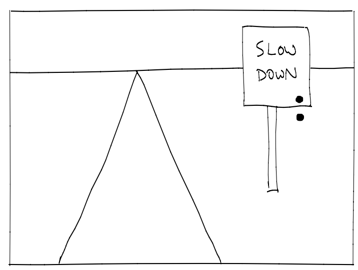
4. Suppose you are a passenger in a car and you wish to read a detailed road sign on the right side of the road as your car approaches the sign and moves past it. Describe how the three types of eye movements (saccades, smooth pursuit, VOR) are used for you to perform this reading task. In particular, what will be the directions of rotation of the eye relative to the head for the three types of eye movements ?

5. When there is a depth discontinuity in the scene, the translation component of the image motion field will have a discontinuity as well, since the translation component of the motion field depends on (inverse) depth.

   [ASIDE: Local differences in the translation component of the motion field which is due to a depth difference is called *motion parallax*. Motion parallax is only produced by the translation component of observer motion. It is not produced by the rotation component. The reason is that the rotation component of the motion field does not depend on the depth Z. Intuitively, the *angular distance* between any two points in the visual field doesn't depend on the direction in which the viewer is looking.]

(a) The image below depicts a scene viewed through a car window, *i.e.* the image boundary is the frame of the window. Two black dots are show on the right. The upper one marks a point on a road sign. The lower one marks a point on the ground.

Sketch the 2D image velocity vectors at the two positions marked by black dots. Assume that you are looking in the direction that the car is driving i.e. the horizon point at the tip of the triangle.



(b) Same question, but now suppose you are making a smooth pursuit eye movement to track the point marked by the upper dot. Draw the *retinal* image velocity vectors at these two marked points. Briefly explain.

Solutions

1. The information given is enough to estimate depth *up to an unknown scale factor*. The idea is that if you know the motion vector (v_x, v_y) in some visual direction (x, y) and you know the direction of heading $(\frac{T_x}{T_z}, \frac{T_y}{T_z})$, then you can use the motion field equation to solve for $Z(x, y)$. The trouble is that you also need to know T_z.

   The basic problem is that the direction of heading $(\frac{T_x}{T_z}, \frac{T_y}{T_z})$ contains only two values, not three. The translation speed of the observer is known up to a missing scale parameter. For example, if you were to double the size of the scene (so each point (X, Y, Z) becomes $(2X, 2Y, 2Z)$ and also also double the speed of the observer $(T_x, T_y, T_z) \rightarrow (2T_x, 2T_y, 2T_z)$ then you would get the same motion field.

2. If the ground plane is at $Y = -h$, i.e. below the origin/eye, then from lecture 1, we have

   $$\frac{y}{f} = -\frac{h}{Z}$$

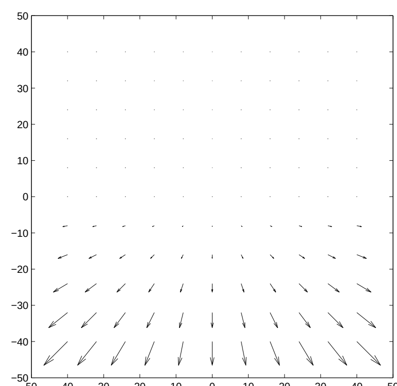
   In lecture 9, we saw that the motion field for forward translation is

   $$(v_x, v_y) = \frac{T_Z}{Z}(\frac{x}{f}, \frac{y}{f})$$

   where (x, y) be horizontal and vertical angles away from the origin. Substituting for $\frac{1}{Z}$, we get

   $$(v_x, v_y) = -\frac{T_Z}{h}(\frac{xy}{f^2}, \frac{y^2}{f^2})$$

   This solution only holds for $y < 0$ since, above the horizon ($y = 0$), the sky is visible and the sky has depth $Z = \infty$ and so the velocity field is 0 there. Below is a sketch of the velocity field. Seem familiar?

3. The trick to this question about the tunnel is to notice that the scene is radially symmetric about the z axis. Thus, if the observer is moving forward, then the velocity field is radially symmetric too. So if we were to parameterize the image plane by polar coordinates (r, θ) instead of (x, y), then the speed would depend on r but not θ.

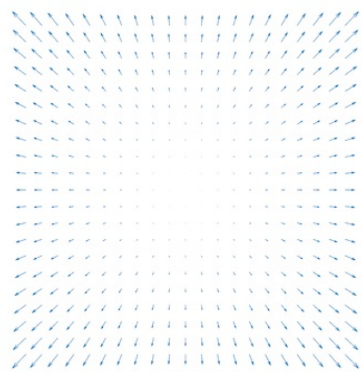
Since the tunnel cylinder is of radius h, we can use the solution from the previous question. Take the plane $X = 0$. The velocity field on the image line $x = 0$ and $y < 0$ is

$$(v_x, v_y) = \frac{T_Z}{h}(0, -\frac{y^2}{f^2}).$$

For a radially symmetric solution, the speed will depend on the distance $r = \sqrt{x^2 + y^2}$ from the origin. Since the speed for line $x = 0$ is y^2/f^2, the speed for general (x, y) must be $(x^2 + y^2)/f^2$. So,

$$(v_x, v_y) = \sqrt{x^2 + y^2}\ \frac{T_Z}{hf^2}(x, y).$$

A *sketch* of the field is shown here. This field is radially expanding, but it is different from the field one sees when moving toward a wall. The main difference is that the speed here goes like $(x^2 + y^2)/f^2$, whereas when one moves toward a wall, the speed varies with $\sqrt{x^2 + y^2}/f$.



4. - The sign is on the right side of the road and so the head may turn to the right as the car drives past, in order to keep the sign close to the center of gaze. When the head turns, the eyes automatically turn left to cancel the retinal motion that is due to head motion (VOR).
   - Regardless of the head rotation and the VOR that cancels it, pursuit eye movements may still be needed to track the sign and reduce retinal velocity to 0. We cannot say what the direction of rotation is because it depends on how much head rotation there. If the head doesn't rotate enough to cancel the motion, the pursuit motions would be to the right. If the head moves too much overcompensates for rotation, the pursuit motions would be to the left.
   - A saccade will be used initially to move the eye to the sign. Further saccades will be need to jump from word to word as you read the sign. (The jumping is from left to right within a line, and then back to the left and down to start a new line.)

5. For (a), the velocities point away from the direction of heading (the end of the road). The two vectors should be close to parallel. However, the speed varies with inverse depth, so the point on the sign has greater speed (longer vector) than the point on the ground.

   For (b), the upper black point has zero retinal velocity since you are tracking it with a pursuit movement which cancels the velocity at that point in (a). For the lower one, you take the lower vector in (a) and subtract roughly the upper vector in (a), since this is the vector that comes from the smooth pursuit.