- Review: The Optical Flow Problem
- Computing the Error of an Estimate
- Computing the Estimate with Minimum Error
- An Aside: Geometry of Matrices

1 Review: The Optical Flow Problem

Last time, the optical flow equation was introduced:

$$\dot{\mathbf{q}} = \lambda \hat{\mathbf{q}}^2 \mathbf{t} + \hat{\mathbf{q}}\omega \tag{1}$$

where $\mathbf{q}$ is the direction, λ is the nearness ($\frac{1}{||\mathbf{x}||}$), $\mathbf{t}$ is the translation, and ω is the rotation. Given a set of N point-velocity pairs

$$\{(\mathbf{q}_1, \dot{\mathbf{q}}_1), (\mathbf{q}_2, \dot{\mathbf{q}}_2), \ldots, (\mathbf{q}_N, \dot{\mathbf{q}}_N)\}$$

(which might be computed from two sets of image points and a correspondence mapping), we would like to ask the question, "What is the most likely translation $\mathbf{t}$, rotation ω, and reduced nearnesses μ_i for the points $\mathbf{q}_i$ which would cause the observed velocities $\dot{\mathbf{q}}_i$?"[1] In the previous lecture, we saw that with five such pairs, the $\mathbf{t}$, ω, and μ_i are exactly determined; however, in the presence of noise, an exact solution is likely to be incorrect. To insure that our estimation of the actual $\mathbf{t}$, ω, and μ_i is good, we can estimate the "best" solution to the overdetermined system arising from more than five pairs. To do this, we will use least-squares optimization.

2 Computing the Error of an Estimate

Define $\mathbf{p}_i$ as $\dot{\mathbf{q}}_i$, the measured parallax, and define $\mathbf{p}'_i$ as the value of $\dot{\mathbf{q}}$ in equation (1), the predicted parallax given an estimate of $\langle \mathbf{t}, \omega, \mu_i \rangle$. Then we can define the error of our estimate of $\langle \mathbf{t}, \omega, \{\mu_i \,|\, 1 \leq i \leq N\} \rangle$ as follows:

$$E = \frac{1}{N} \sum_{i=1}^{N} ||\mathbf{p}_i - \mathbf{p}'_i||^2 \tag{2}$$

[1] Recall that the reduced nearness μ_i is the value of the nearness λ_i when $\mathbf{t}$ is normalized.

3 Computing the Estimate with Minimum Error

Since E is a function of our estimate $\langle \mathbf{t}, \omega, \{\mu_i \,|\, 1 \leq i \leq N\}\rangle$, we can find its global minima by locating the points where its gradient is zero. To do this, we must find the partial derivatives of E with respect to our parameters:

$$\frac{\partial E}{\partial \mathbf{t}} = \frac{1}{N}\sum_{i=1}^{N} 2\mu_i \widehat{\mathbf{d}}_i^2(\mathbf{p}_i - \mathbf{p}_i') \tag{3}$$

$$\frac{\partial E}{\partial \mu_i} = 2(\widehat{\mathbf{d}}_i^2\mathbf{t})^T(\mathbf{p}_i - \mathbf{p}_i') \tag{4}$$

$$\frac{\partial E}{\partial \omega} = \frac{1}{N}\sum_{i=1}^{N} 2\widehat{\mathbf{d}}_i(\mathbf{p}_i - \mathbf{p}_i') \tag{5}$$

We want to find an estimate of $\langle \mathbf{t}, \omega, \{\mu_i \,|\, 1 \leq i \leq N\}\rangle$ such that

$$\frac{\partial E}{\partial \mathbf{t}} = \frac{\partial E}{\partial \mu_i} = \frac{\partial E}{\partial \omega} = 0$$

where $\mathbf{d}_i = \mathbf{q}_i$. We must be careful to add the constraint that $||\mathbf{t}|| = 1$ (since we are using reduced nearness values).

Now, all the $\mathbf{d}_i$ and $\mathbf{p}_i$ are known; they are the $(\mathbf{q}_i, \dot{\mathbf{q}}_i)$ pairs given to us. The $\mathbf{p}_i'$ are unknown, as in them are the hidden unknowns $\mathbf{t}$, ω, and μ_i. Equations (3) and (4) are nonlinear in $\mathbf{d}_i$, so linear optimization techniques cannot be applied. Instead we will use a nonlinear iterative technique; we will begin with a guess for $\langle \mathbf{t}, \omega, \{\mu_i \,|\, 1 \leq i \leq N\}\rangle$ and iterate until the error E stays within some preset tolerance.

3.1 The Iteration Step

Substituting the RHS of equation (1) for $\mathbf{p}_i'$ in equation (3) and setting the partial derivative equal to zero, we get

$$\frac{1}{N}\sum_{i=1}^{N}\left[-\mu_i\mathbf{p}_i - \mu_i(\mu_i\mathbf{t} - \mu_i(\mathbf{d}_i^T\mathbf{t})\mathbf{d}_i + \widehat{\omega}\mathbf{d}_i\right] = 0$$

Rearranging the terms, we have

$$\frac{1}{N}\sum_{i=1}^{N} -\mu_i^2\mathbf{t} = -\frac{1}{N}\sum_{i=1}^{N}\mu_i\mathbf{p}_i + \frac{1}{N}\sum_{i=1}^{N}\mu_i^2(\mathbf{d}_i^T\mathbf{t})\mathbf{d}_i - \frac{1}{N}\sum_{i=1}^{N}\widehat{\omega}\mu_i\mathbf{d}_i$$

Recognizing each of the summations as an average over all N pairs, we can rewrite this as

$$\left\langle \mu^2 \right\rangle \mathbf{t} = -\langle \mu\mathbf{p}\rangle + \left\langle \mu^2(\mathbf{d}^T\mathbf{t})\mathbf{d}\right\rangle - \widehat{\omega}\langle \mu\mathbf{d}\rangle$$

where $\langle \alpha_i \rangle$ is shorthand for the average of all α_i, $1 \leq i \leq N$. If we let $\xi = 1/\langle \mu^2 \rangle$, then we have an equation which can be used to "predict" a new value of $\mathbf{t}$, given a current estimate:

$$\mathbf{t}_{\text{new}} = \xi\left(-\langle \mu\mathbf{p}\rangle + \left\langle \mu^2(\mathbf{d}^T\mathbf{t}_{\text{old}})\mathbf{d}\right\rangle - \widehat{\omega}\langle \mu\mathbf{d}\rangle\right) \tag{6}$$

In a similar fashion, we can derive predictors of ω and the μ_i:

$$\mu_{\mathbf{i}_{\text{new}}} = \frac{-\mathbf{t}^T(\mathbf{p}_i + \widehat{\omega}\mathbf{d}_i)}{1 - (\mathbf{d}_i^T\mathbf{t})^2} \tag{7}$$

$$\omega_{\text{new}} = \frac{\langle \widehat{\mathbf{p}}\mathbf{d}\rangle + \widehat{\mathbf{t}}\left\langle \mu\widehat{\mathbf{d}}\right\rangle}{\langle I - \mathbf{d}\mathbf{d}^T\rangle} \tag{8}$$

Equations (6), (7), and (8) give us the means to obtain better and better approximations to the estimate $\langle \mathbf{t}, \omega, \{\mu_i \,|\, 1 \le i \le N\}\rangle$ which minimizes our error function E. Given some initial estimate, we can evaluate each of equations (6), (7), and (8) to obtain a better estimate, repeating the process until the error function decreases by less than some preset tolerance. The only remaining question is how to obtain the initial estimate to start the procedure.

3.2 Initialization

First we observe that once we fix $\mathbf{t}$ and ω, all of the μ_i are determined by equation (1). One approach to estimating $\mathbf{t}$ and ω is to pick five of the point-velocity pairs and solve explicitly for $\mathbf{t}$ and ω. However, the error function E may have local minima, and it is possible this method will not find the global minima. In practice, it is better to choose $\mathbf{t}$ and ω by assuming one of the two is zero; i.e., estimate the translation component from the data as if there were no rotation, and vice-versa for the rotational component.

To initialize ω in this way, we can simply set $\mathbf{t} = \mathbf{0}$ in equation (8) and solve for ω. This is possible since the predicted value of ω does not rely on a previous prediction of its value. However, the predicted value of $\mathbf{t}$ in equation (6) does depend upon a previous prediction of $\mathbf{t}$; so we must find another way to initialize the value of $\mathbf{t}$.

Note that when $\omega = \mathbf{0}$, the parallax vectors $\mathbf{p}_i$ arise purely from translation. Consider a single parallax vector $\mathbf{p}_i$ on the tangent plane to the viewing sphere, and its corresponding distance vector $\mathbf{d}_i$. As we are assuming there is no rotational component to the parallax $\mathbf{p}_i$, it must be true that $\mathbf{t}$, $\mathbf{p}_i$ and $\mathbf{d}_i$ all lie in a single plane. Thus, if we contruct $\mathbf{v}_i$ as a normal to this plane, i.e., $\mathbf{v}_i = \mathbf{p}_i \wedge \mathbf{d}_i$, then $\mathbf{t}$ will have no projection on $\mathbf{v}_i$, i.e., $\mathbf{v}_i \wedge \mathbf{t} = 0$. Since this will only be true when $\omega = \mathbf{0}$, we can use

$$E_{\mathbf{t}} = \sum_{i=1}^{N} (\mathbf{t} \cdot \mathbf{v}_i)^2$$

as an error function, and minimize it. Manipulating this equation, we see

$$\begin{aligned} E_{\mathbf{t}} &= \sum_{i=1}^{N} \left(\mathbf{t}^T\mathbf{v}_i\right)\left(\mathbf{v}_i^T\mathbf{t}\right) \\ &= \sum_{i=1}^{N} \mathbf{t}^T\left(\mathbf{v}_i\mathbf{v}_i^T\right)\mathbf{t} \\ &= \mathbf{t}^T\left(\sum_{i=1}^{N}\mathbf{v}_i\mathbf{v}_i^T\right)\mathbf{t} \end{aligned}$$

The $\mathbf{t}$ which minimizes $E_{\mathbf{t}}$ is the eigenvector of $\sum_{i=1}^{N} \mathbf{v}_i \mathbf{v}_i^T$ with the smallest eigenvalue! Thus, to compute our initial estimate of $\mathbf{t}$, we can form the matrix $\sum_{i=1}^{N} \mathbf{v}_i \mathbf{v}_i^T$ and find the eigenvector with smallest eigenvalue. (Note that this matrix is symmetric positive-definite, and so all its eigenvalues will be real.)

4 An Aside: The Geometry of Matrices

It is often useful to visualize the geometric transforms represented by two- and three-dimensional matrices. This can most easily be accomplished by using a *gauge figure*, in which a familiar shape, such as the unit circle, is subjected to the transform to illuminate the transform's effects.

For example, the identity matrix I_2 maps the unit circle to itself (as one would expect). Scalings of the identity matrix αI_2 are isotropic magnifications, e.g., the matrix

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

magnifies the unit circle to a circle of radius two. The matrix

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$$

is an anisotropic magnification, since it expands the y-axis more than the x-axis. The matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \cos(\pi/2) & -\sin(\pi/2) \\ \sin(\pi/2) & \cos(\pi/2) \end{bmatrix}$$

rotates the xy-plane by 90 degrees, since

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

is a rotation by θ.