CS 280, Computer Vision (Prof. Jitendra Malik) Feb 10, 1999
Lecture #7 (Optical Flow (Continued)) **DRAFT** Notes by Stephen David

- Spherical Projection of Optical Flow
  - Focus of expansion
  - Scale factor ambiguity
  - Parallax
- Inverse Problem: Motion from Flow Field
  - Least squared error fitting
  - Error for motion estimation
  - Gradient of a function

1 Spherical Projection of Optical Flow

In the previous lecture, we derived an equation for the spherical projection of optical flow for a viewer moving with translation $\vec{t}$ and rotation $\vec{\omega}$:

$$\dot{q} = -\frac{1}{||\vec{X}||}(\vec{t} - (\vec{t}.\,\vec{q})\vec{q}) - \vec{\omega} \wedge \vec{q} \tag{1}$$

where $\vec{x}$ is the position of a point in space and $\vec{q}$ is its projection onto the viewing sphere.

We can write this equation in a simpler way if we make some substitutions. Recall from earlier lectures that we defined $\hat{q}$ to be the skew-symmetric matrix which calculates cross product by $\vec{q}$ (*i.e.,* $\hat{q}\vec{x} = \vec{q} \wedge \vec{x}$). If we let $||\vec{q}|| = 1$, then we also found that

$$\hat{q}^2 = \vec{q}\vec{q}^T - I.$$

Additionally, we define **nearness**, λ, to be

$$\lambda = \frac{1}{||\vec{X}||} = \frac{1}{r},$$

where r is the distance to the point $\vec{X}$.

Substituting these values into equation **??**, we get:

$$\dot{q} = \lambda\hat{q}^2\vec{t} + \hat{q}\vec{\omega}. \tag{2}$$

1.1 Focus of expansion

Suppose we observe the special case of a system undergoing pure translation ($\vec{\omega} = 0$). In that case,

$$\dot{q} = \lambda \hat{q}^2 \vec{t} = \lambda \vec{q} \wedge (\vec{q} \wedge \vec{t}).$$

Notice that $\dot{q} = 0$ when $\vec{q} || \vec{t}$. This value of $\vec{q}$ is called the **focus of expansion**, and it indicates the direction of pure translational motion.

Suppose the camera is undergoing pure rotation ($\vec{t} = 0$). In that case, $\dot{q} = \vec{q} \wedge \vec{\omega}$, and $\dot{q} = 0$ at points parallel to the access of rotation (*i.e.*, $\vec{q} || \vec{\omega}$).

1.2 Scale factor ambiguity

Because projection from a 3-d world onto a 2-d view sphere necessarily loses some information about the scene, certain features of the scene can not be recovered from optical flow alone. More specifically, we can measure the product, $\lambda ||\vec{t}||$, but we can not measure the ratio of λ to $||\vec{t}||$. This effect, **scale factor ambiguity**, gives rise to the phenomenon that objects moving slowly at a short distance appear similar to object far away moving in the same direction at a higher speed.

1.3 Parallax

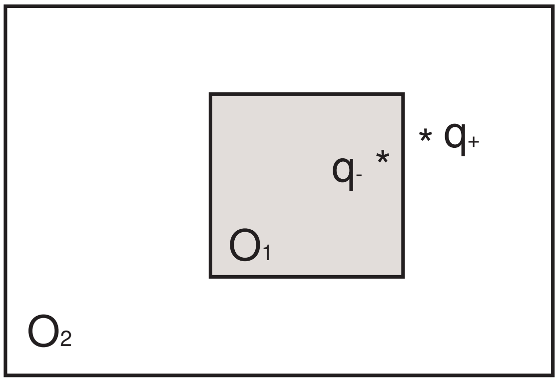

Figure 1: Example of parallax. Points near the boundary of O_1 and O_2 show different optical flows because they lie at different distances.

Optical flow provides useful information about depth differences between two adjacent objects in a scene. Let $\vec{q}_-$ and $\vec{q}_+$ be two points in the image near the edge of two objects, O_1 and O_2 (see figure **??**), at different depths, λ_- and λ_+ in a scene. The corresponding optical flow for each of these points are

$$\dot{q}_- = \lambda_- \hat{q}_-^2 \vec{t} + \hat{q}_- \vec{\omega}.$$

and

$$\dot{q}_+ = \lambda_+ \hat{q}_+^2 \vec{t} + \hat{q}_+ \vec{\omega}$$

At points near the boundary of O_1 and O_2 in the image, $\vec{q} = \vec{q}_- \approx \vec{q}_+$. We define **parallax** as the difference between the two flows:

$$\vec{q}_- - \vec{q}_+ = (\lambda_- - \lambda_+)\hat{q}^2\vec{t} + 0$$

The rotational component of flow cancels out, but the translational component remains, due to the different values of λ near the edge.

2 Inverse Problem: Motion from Flow Field

Suppose we can measure optical flow at a number of points in the visual field. (This assumes that we have solved the correspondence problem required to identify sets of two points in two different images as the same, but that's another issue reserved for later in the semester.) We wish to use optical flow data to calculate how a scene is translating and rotating. Referring to equation **??**, we have values for $\vec{q}_i$ and $\dot{q}_i$ where $i = 1 \ldots N$, and we would like to find values for $\vec{t}$, $\vec{\omega}$, and λ_i. Note that the values of $\vec{t}$ and $\vec{\omega}$ are constant for the entire image, but λ_i can change with each point $\vec{q}_i$.

Since $\vec{t}$ and $\vec{\omega}$ are both 3-d vectors and each λ_i is a scalar, we are looking for $6 + N$ unknowns. Due to the scale factor ambiguity, however, we can only hope to resolve $5 + N$ of those unknowns. Each $\vec{q}_i$ is a 2-d vector, giving two equations. In theory, we need a minumum of $N = 5$ points to give us 10 equations and 10 unknowns to solve the problem exactly. Because of noise, however, it's generally better to have a value of N significantly larger than 5.

2.1 Least squared error fitting

We solve the motion from flow problem using a least square error method. The idea is that because there is noise in the system, we aren't likely to find a solution that would give us the same values of flow back if we were to plug $\vec{q}_i$ and our calculated values for $\vec{t}$, $\vec{\omega}$, and λ_i back into equation **??**. Instead, we try to minimize the square of the error between our approximation of flow, $\dot{q}_i'$ and the actual value of $\dot{q}_i$.

How do we measure error? Suppose we have a some function $y = f(x)$ for which we've measured N values of (x_i, y_i). We approximate f as a linear function,

$$f(x) = a_1 x + a_2.$$

The error of our approximation is:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - a_1 x_i - a_2)^2 \tag{3}$$

2.2 Error for motion estimation

We use the notation of Koenderink & Van Doorn ("Facts on Optic Flow", *Biol. Cybernetics.* 56, 247-254, 1987) for our MSE calculation. Let $d_i = q_i$, $p_i = \dot{q}_i$, and $\mu_i = \lambda_i$. Furthermore, assume $||\vec{t}|| = 1$ to account for scale factor ambiguity. We continue to use $\vec{t}$ and $\vec{\omega}$ for our motion vectors.

We use equation **??** to calculate error:

$$E = \frac{1}{N}\sum_{i=1}^{N} ||\vec{p}_i - \vec{p}_i'||^2 \tag{4}$$

Where we rewrite equation **??**: $\vec{p}_i' = \mu_i \hat{d}_i^{\,2}\vec{t} + \hat{d}_i\vec{\omega}$.

2.3 Gradient of a function

Now that we've got the MSE, how do we find the $\vec{t}$, $\vec{\omega}$, and μ_i to minimize it? As we learned in calculus, we know that a function will reach its local minima and maxima at points where the derivative is 0.

To find the derivative of the MSE with respect to its vector parameters, we must take the **gradient**. Recall the definition of gradient for $f : \Re^3 \to \Re$:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix}.$$

So, for example, if $f = 2x + 3xy + z^2$,

$$\nabla f = \begin{bmatrix} 2 \\ 2x \\ 2z \end{bmatrix}.$$

For $f = \vec{y}^T\vec{x} = \vec{y} \wedge \vec{x}$,,

$$\nabla f = \begin{bmatrix} y \\ x \end{bmatrix}.$$

And for $f = ||\vec{x}||^2$,

$$\nabla f = 2\vec{x}.$$

We take the derivative of E in equation /refflowMSE with respect to each of its three paramters, and use some of the properties of the equations above:

$$\frac{\partial E}{\partial \vec{t}} = -\frac{1}{N}\sum_{i=1}^{N} 2(\vec{p}_i - \vec{p}_i')\mu_i\hat{d}_i^2 = 0 = (3 \times 1 \text{ vector})$$

$$\frac{\partial E}{\partial \mu_i} = (\hat{d}_i^2\vec{t})^T(\vec{p}_i - \vec{p}_i') = 0 = (1 \times 1 \text{ scalar})$$

$$\frac{\partial E}{\partial \vec{\omega}} = -\frac{1}{N}\sum_{i=1}^{N} 2(\vec{p}_i - \vec{p}_i')\hat{d}_i = 0 = (3 \times 1 \text{ vector})$$

Useful identities:

- $\hat{d}_i^2 \vec{p}_i = \vec{p}_i$
- $\hat{d}_i^2 \vec{p}_i' = \mu_i \vec{t}(\vec{d}_i^T \vec{t})\vec{d}_i + \hat{\omega}\vec{d}_i$

Use these identities to rewrite our error with respect to $\vec{t}$:

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

And rearrange to get:

$$\left(\frac{1}{N}\sum_{i=1}^{N}\mu_i^2\right)\vec{t} = - < \mu\vec{p} > + < \mu^2(\vec{d}^T\vec{t})d > - \hat{\omega} < \mu\vec{d} >$$

Note: $< \ldots >$ indicates expected value of the term inside.
To be continued in next lecture ...