JOHNS HOPKINS UNIVERSITY

Motion

How to measure image velocity or optical flow

Topics: (1) Short-range motion

(2) Long-range motion

(3) Radial bases motion

Motion

The task is the estimate the correspondence between image pixels in a sequence of time frames

Most theories address the special case with only two time frames

$$I(\mathbf{x},t), I(\mathbf{x},t+\Delta t)$$

This is an example of a correspondence problem like binocular stereo, and we will discuss later in the course.

Shot-range correspondence differs from binocular stereo because

(i) There is no epipolar line constraint, so the correspondence problem is 2-dimensional

(ii) The motion or displacement between pixels in neighboring time frames is usually small

Simplest formulation

Match points with similar intensity values

$$E_1\left[\{\mathbf{v}(\mathbf{x})\}\right] = \int \{I(\mathbf{x}, t+\Delta t) - I(\mathbf{x}+\mathbf{v}(\mathbf{x}), t\}^2 \, d\mathbf{x}$$

where $\mathbf{v}(\mathbf{x})$ is the motion, or displacement

This formulation is ill-posed, there are many possible solutions $\{\hat{\mathbf{v}}(\mathbf{x})\} = \arg\min E[\mathbf{v}(\mathbf{x})]$

To regularize the problem, we can use the prior assumption that motion is likely to be smooth

$$E_2\left[\{\mathbf{v}(\mathbf{x})\}\right] = \int \{I(\mathbf{x}, t+\Delta t) - I(\mathbf{x}+\mathbf{v}(\mathbf{x}), t\}^2 \, d\mathbf{x} + \lambda \int \{\nabla \mathbf{v}(\mathbf{x}) \cdot \nabla \mathbf{v}(\mathbf{x})\} \, d\mathbf{x}$$

$E_2[\cdot]$ is a non-convex function of $\{\mathbf{v}(\mathbf{x})\}$, so there is no natural algorithm to minimize it and solve for $\{\hat{\mathbf{v}}(\mathbf{x})\} = \arg\min E_2[\{\mathbf{v}(\mathbf{x})\}]$

Instead, Horn & Schunk suggested replacing the first term by exploiting the fact that $\mathbf{v}(\mathbf{x})$ is small, and deriving the optical flow constraint:

$$\mathbf{v} \cdot \nabla I + \frac{\partial I}{\partial t} = 0$$

This follows by assuming that the motion is locally constant so that

$I(\mathbf{x}, t) = F(\mathbf{x} - \mathbf{v}t)$, for some function $\mathrm{F}(\cdot)$

$$\Rightarrow \quad \nabla I = \nabla F, \frac{\partial I}{\partial t} = -\mathbf{v} \cdot \nabla F \quad \Rightarrow \quad v \cdot \nabla I + \frac{\partial I}{\partial t} = 0$$

This yields a convex energy function: Horn & Schunk's formulation

$$E_3\left[\{\mathbf{v}(\mathbf{x})\}\right] = \int \left\{ \mathbf{v}(x) \bullet \nabla I(\mathbf{x}) + \frac{\partial I}{\partial \mathbf{x}} \right\}^2 d\mathbf{x} + \lambda \int \{\nabla \mathbf{v}(\mathbf{x}) \cdot \nabla \mathbf{v}(\mathbf{x})\} \, d\mathbf{x}$$

It can be minimized by steepest descent

Horn & Schunk's formulation

There are many variations of the approach.

- The prior can be changed
- The term $\left\{ \mathbf{v}(\mathbf{x}) \cdot \nabla I(\mathbf{x}) + \frac{\partial I}{\partial \mathbf{x}} \right\}$ can be normalized

Taylor series interpretation of optical flow

$$I(\mathbf{x}+\mathbf{v}(\mathbf{x}), t+\Delta t) = I(\mathbf{x}, t) + \mathbf{v} \cdot \nabla I + \Delta t \frac{\partial I}{\partial t}$$

$$I(\mathbf{x}+\mathbf{v}(\mathbf{x}), t+\Delta t) - I(\mathbf{x}, t) = \mathbf{v} \cdot \nabla I + \Delta t \frac{\partial I}{\partial t}$$

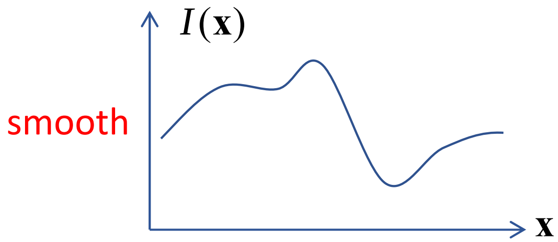
The Taylor series interpretation allows us to derive coarse-to-fine approach which addresses a problem in the optical flow equation

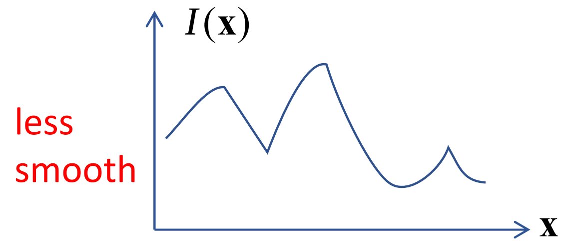
Optical flow problem: $\mathbf{v}\cdot\nabla I + \frac{\partial I}{\partial t} = 0$

Must be discretized when we calculate it as an image lattice

The derivation must be approximated by difference (e.g. $\frac{dI}{d\mathbf{x}} \approx \frac{I(\mathbf{x}+\Delta) - I(\mathbf{x})}{\Delta}$)

This is problematic unless the image is smooth





Strategy: (1) Smoothen the image by a Gaussian $G_\sigma * I(\mathbf{x})$ estimate the optical flow $\mathbf{v}_1(\mathbf{x})$

(2) Do a Taylor series expansion about $\mathbf{x}+\mathbf{v}_1(\mathbf{x})$ for small displacement $\mathbf{v}_2(\mathbf{x})$

JOHNS HOPKINS UNIVERSITY

Taylor series expansion

$$I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x})+\mathbf{v}_2(\mathbf{x}),t+\Delta t\right)=I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x}),t\right)+\mathbf{v}_2(\mathbf{x})\nabla I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x}),t\right)+\Delta t\frac{\partial}{\partial t}I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x})\right)$$

$\Rightarrow$ $$I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x})+\mathbf{v}_2(\mathbf{x}),t+\Delta t\right)-I(\mathbf{x},t)$$

$$=I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x}),t\right)-I(\mathbf{x},t)+\mathbf{v}_2(\mathbf{x})\cdot\nabla I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x}),t\right)+\Delta t\frac{\partial}{\partial t}I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x})\right)$$

Replace $\left\{I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x})+\mathbf{v}_2(\mathbf{x}),t+\Delta t\right)-I(\mathbf{x},t)\right\}^2$

by $\left\{I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x}),t\right)-I(\mathbf{x},t)+\mathbf{v}_2(\mathbf{x})\cdot\nabla I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x}),t\right)+\Delta t\frac{\partial}{\partial t}I\left(\mathbf{x}+\mathbf{v}_1(\mathbf{x})\right)\right\}^2$

Key idea: Series of approximations at different smoothed image

More Advanced Models

(1) Replace intensity $I(\mathbf{x}, \mathrm{t})$ by image features

(2) Learn the correspondence terms using Machine Learning, e.g., Deep Networks

(3) Modify the smoothness term

(4) Add motion layers to deal with transparent motion, or objects moving in front of moving background

(5) Dealing with occlusions, points that are unmatched

(6) Motion over true sequences, Kalman Filter (later in the course)

(7) 2D motion to 3D motion