MA3113: Topics in Mathematical Image Processing I
Optical Flow Estimation

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: March 01, 2022/Last updated: April 1, 2026

Optical flow estimation

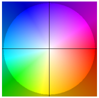
Let $I : \Omega \times [0, T] \subset \mathbb{R}^3 \to \mathbb{R}$ be a given image sequence, of grayscale values in space $(x, y) \in \Omega$ and time $t \in [0, T]$.

- The optical flow is defined as the velocity field

$$\boldsymbol{w}(x, y) = (u(x, y), v(x, y))^\top,$$

  such that one frame $I(x, y, t)$ is translated to the next frame $I(x + \delta x, y + \delta y, t + \delta t)$ by the mapping $(\delta x, \delta y, \delta t) = (u\delta t, v\delta t, \delta t)$.

- Optical flow estimation aims to compute the velocity field $(u, v)^\top$ from two consecutive images.



https://learnopencv.com/optical-flow-using-deep-learning-raft/

The brightness constancy assumption

- The brightness constancy assumption (BCA) is a fundamental principle in optical flow estimation, stating that the intensity of a moving object remains unchanged between consecutive image frames. *The brightness constancy assumption* can be written as

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t).$$

- Using the first-order Taylor expansion at (x, y, t), we obtain

$$\begin{aligned} I(x + \delta x, y + \delta y, t + \delta t) &= I(x, y, t) + I_x(x, y, t)\delta x + I_y(x, y, t)\delta y \\ &\quad + I_t(x, y, t)\delta t + \mathcal{O}(\delta x^2 + \delta y^2 + \delta t^2), \end{aligned}$$

where $\nabla I := (I_x, I_y)^\top$ and I_t denote spatial and temporal partial derivatives of the image I, respectively. With the BCA, we obtain

$$I_x(x, y, t)\delta x + I_y(x, y, t)\delta y + I_t(x, y, t)\delta t + \mathcal{O}(\delta x^2 + \delta y^2 + \delta t^2) = 0.$$

Dividing by δt and letting $\delta t \to 0^+$, we have

$$\nabla I \cdot \boldsymbol{w} + I_t = 0.$$

The derivation of the BCA from another point of view

From another point of view, the brightness constancy assumption can be represented as

$$I(x(t), y(t), t) = c.$$

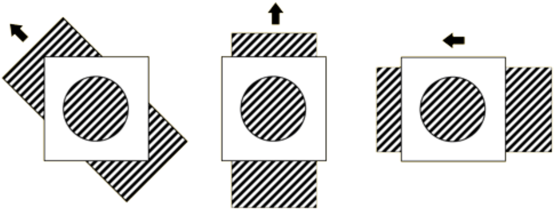
Taking derivatives of I with respect to t by the chain rule, we obtain

$$\begin{aligned} 0 &= \frac{d}{dt} I(x(t), y(t), t) \\ &= I_x \frac{dx}{dt} + I_y \frac{dy}{dt} + I_t \frac{dt}{dt} \\ &= I_x u + I_y v + I_t \\ &= \nabla I \cdot \boldsymbol{w} + I_t. \end{aligned}$$

The aperture problem

The aperture problem is a fundamental ambiguity in visual perception and computer vision in which the local motion of an edge viewed through a small aperture cannot be uniquely determined.

The aperture problem often occurs in methods that rely on local pixel information to estimate optical flow.



In the field of optical flow estimation, the two most classic models are the *Lucas-Kanade model* and the *Horn-Schunck model.* The former uses local information, and the latter employs global information.

The Lucas-Kanade model

Lucas and Kanade proposed a local smoothing method for estimating the optical flow. Given two consecutive images at times t_k and t_{k+1} and fixed a position $\boldsymbol{z}_0$, the optical flow $(u(\boldsymbol{z}_0), v(\boldsymbol{z}_0))^\top$ at $\boldsymbol{z}_0$ is given by solving the following optimization problem:

$$\min_{u,v} E_{LK}(u,v) := \sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} \Big(I_x(\boldsymbol{p}_r, t_k)u + I_y(\boldsymbol{p}_r, t_k)v + I_t(\boldsymbol{p}_r, t_k) \Big)^2.$$

This method assumes that pixels $\boldsymbol{p}_r$, $r = 1, 2, \cdots, n$, in the small window $\mathcal{W}(\boldsymbol{z}_0)$ center at z_0 share the same velocity $(u(\boldsymbol{z}_0), v(\boldsymbol{z}_0))^\top$. Denote

$$\boldsymbol{A} = \begin{bmatrix} I_x(\boldsymbol{p}_1, t_k) & I_y(\boldsymbol{p}_1, t_k) \\ I_x(\boldsymbol{p}_2, t_k) & I_y(\boldsymbol{p}_2, t_k) \\ \vdots & \vdots \\ I_x(\boldsymbol{p}_n, t_k) & I_y(\boldsymbol{p}_n, t_k) \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} -I_t(\boldsymbol{p}_1, t_k) \\ -I_t(\boldsymbol{p}_2, t_k) \\ \vdots \\ -I_t(\boldsymbol{p}_n, t_k) \end{bmatrix}.$$

We can drop the third variable t_k when there is no risk for confusion.

The least squares problem

It comes to a least squares problem:

$$\min_{u,v} \|Aw - b\|_2^2,$$

and the normal equation is given by

$$A^\top A w = A^\top b.$$

Suppose rank $(A) = 2$. Then $A^\top A$ is nonsingular and the least square solution is given by

$$\begin{aligned} w &= (A^\top A)^{-1} A^\top b \\ &= \begin{bmatrix} \sum_{p_r \in \mathcal{W}(z_0)} I_x^2(p_r) & \sum_{p_r \in \mathcal{W}(z_0)} I_x(p_r) I_y(p_r) \\ \sum_{p_r \in \mathcal{W}(z_0)} I_x(p_r) I_y(p_r) & \sum_{p_r \in \mathcal{W}(z_0)} I_y^2(p_r) \end{bmatrix}^{-1} \begin{bmatrix} \sum_{p_r \in \mathcal{W}(z_0)} -I_x(p_r) I_t(p_r) \\ \sum_{p_r \in \mathcal{W}(z_0)} -I_y(p_r) I_t(p_r) \end{bmatrix}. \end{aligned}$$

Finite difference approximations

Below, we will apply forward differences to the spatial and temporal derivatives in the normal equation. By Taylor's expansion, we have

$$I_x(x_i, y_j, t_k) = \frac{1}{h}\left(I(x_i + h, y_j, t_k) - I(x_i, y_j, t_k)\right) - \frac{1}{2} I_{xx}(\xi_i, y_j, t_k) h,$$
$$\text{for some } \xi_i \in (x_i, x_i + h).$$

Similarly, applying Taylor's expansion to $I_y(x_i, y_j, t_k)$ and $I_t(x_i, y_j, t_k)$, we obtain the following first-order finite difference approximations:

$$I_x(x_i, y_j, t_k) \approx I_{x,i,j,k} := \frac{1}{h}(I_{i+1,j,k} - I_{i,j,k}),$$
$$I_y(x_i, y_j, t_k) \approx I_{y,i,j,k} := \frac{1}{h}(I_{i,j+1,k} - I_{i,j,k}),$$
$$I_t(x_i, y_j, t_k) \approx I_{t,i,j,k} := \frac{1}{h}(I_{i,j,k+1} - I_{i,j,k}).$$

$A^\top A$ in the normal equation

For $A^\top A$, we have

$$A^\top A = \begin{bmatrix} \displaystyle\sum_{p_r \in \mathcal{W}(z_0)} I_x^2(p_r) & \displaystyle\sum_{p_r \in \mathcal{W}(z_0)} I_x(p_r) I_y(p_r) \\ \displaystyle\sum_{p_r \in \mathcal{W}(z_0)} I_x(p_r) I_y(p_r) & \displaystyle\sum_{p_r \in \mathcal{W}(z_0)} I_y^2(p_r) \end{bmatrix}.$$

The entries in $A^\top A$ can be approximated by

$$\sum_{p_r \in \mathcal{W}(z_0)} I_x^2(p_r) \approx \frac{1}{h^2} \sum_{p_r \in \mathcal{W}(z_0)} \left(I_{i+1,j,k} - I_{i,j,k}\right)^2,$$

$$\sum_{p_r \in \mathcal{W}(z_0)} I_y^2(p_r) \approx \frac{1}{h^2} \sum_{p_r \in \mathcal{W}(z_0)} \left(I_{i,j+1,k} - I_{i,j,k}\right)^2,$$

$$\sum_{p_r \in \mathcal{W}(z_0)} I_x(p_r) I_y(p_r) \approx \frac{1}{h^2} \sum_{p_r \in \mathcal{W}(z_0)} \left(I_{i+1,j,k} - I_{i,j,k}\right)\left(I_{i,j+1,k} - I_{i,j,k}\right).$$

$A^\top b$ in the normal equation

For $\boldsymbol{A}^\top \boldsymbol{b}$, we have

$$\boldsymbol{A}^\top \boldsymbol{b} = \begin{bmatrix} \sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} -I_x(\boldsymbol{p}_r) I_t(\boldsymbol{p}_r) \\ \sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} -I_y(\boldsymbol{p}_r) I_t(\boldsymbol{p}_r) \end{bmatrix}.$$

The components in $\boldsymbol{A}^\top \boldsymbol{b}$ can be approximated by

$$\sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} -I_x(\boldsymbol{p}_r) I_t(\boldsymbol{p}_r) \approx -\frac{1}{h^2} \sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} \left(I_{i+1,j,k} - I_{i,j,k} \right) \left(I_{i,j,k+1} - I_{i,j,k} \right),$$

$$\sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} -I_y(\boldsymbol{p}_r) I_t(\boldsymbol{p}_r) \approx -\frac{1}{h^2} \sum_{\boldsymbol{p}_r \in \mathcal{W}(\boldsymbol{z}_0)} \left(I_{i,j+1,k} - I_{i,j,k} \right) \left(I_{i,j,k+1} - I_{i,j,k} \right).$$

After discretization, the grid-spacing terms $\frac{1}{h^2}$ on both sides of the normal equation cancel out.

The Horn-Schunck model

- In 1981, Horn and Schunck proposed a global smoothing approach by minimizing the following energy functional

$$\mathcal{E}_{HS}(\boldsymbol{w}) = \frac{1}{2}\int_{\Omega} (\nabla I \cdot \boldsymbol{w} + I_t)^2 + \lambda(|\nabla u|^2 + |\nabla v|^2)\, dxdy,$$

  where $\lambda > 0$ is the regularization parameter.

- By the calculus of variation, we obtain the Euler-Lagrange equations,

$$\begin{aligned} I_x(I_x u + I_y v + I_t) - \lambda \Delta u &= 0 \quad \text{in } \Omega, \quad (\star_1) \\ I_y(I_x u + I_y v + I_t) - \lambda \Delta v &= 0 \quad \text{in } \Omega, \quad (\star_2) \end{aligned}$$

  with the homogeneous Neumann boundary conditions:

$$\frac{\partial u}{\partial \boldsymbol{n}} = 0 \quad \text{and} \quad \frac{\partial v}{\partial \boldsymbol{n}} = 0 \quad \text{on } \partial\Omega,$$

  and then solve the equations approximately.

- Notice that we are supposedly given two consecutive images at times t_k and t_{k+1}.

The Euler-Lagrange equations and BCs

- Consider the Horn-Schunck energy functional,

$$\mathcal{E}_{HS}(\boldsymbol{w}) = \frac{1}{2}\int_{\Omega} (\nabla I \cdot \boldsymbol{w} + I_t)^2 + \lambda(|\nabla u|^2 + |\nabla v|^2)\, dxdy.$$

- By the calculus of variation,

$$\min \int_{\Omega} L\, d\boldsymbol{x} \;\Rightarrow\; \frac{\partial L}{\partial u} - \nabla \cdot \left(\frac{\partial L}{\partial u_x}, \frac{\partial L}{\partial u_y}\right)^{\top} = 0 \quad \text{in } \Omega, \frac{\partial u}{\partial \boldsymbol{n}} = 0 \quad \text{on } \partial\Omega,$$

we have

$$\begin{aligned} (\nabla I \cdot \boldsymbol{w} + I_t)I_x - \nabla \cdot (\lambda u_x, \lambda u_y)^{\top} &= 0 \quad \text{in } \Omega, \\ \frac{\partial u}{\partial \boldsymbol{n}} &= 0 \quad \text{on } \partial\Omega, \\ (\nabla I \cdot \boldsymbol{w} + I_t)I_y - \nabla \cdot (\lambda v_x, \lambda v_y)^{\top} &= 0 \quad \text{in } \Omega, \\ \frac{\partial v}{\partial \boldsymbol{n}} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Finite difference approximations

First, $\Delta u(x_i, y_j)$ can be discretized as follows:

$$\Delta u(x_i, y_j) \approx \frac{1}{h^2}(\bar{u}_{i,j} - 4u_{i,j}), \quad \bar{u}_{i,j} := u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}.$$

Similar approximation can be applied to $\Delta v(x_i, y_j)$. Next, we define

$$\tilde{I}_{x,i,j,k} := I_{i+1,j,k} - I_{i,j,k}, \quad \tilde{I}_{y,i,j,k} := I_{i,j+1,k} - I_{i,j,k}, \quad \tilde{I}_{t,i,j,k} := I_{i,j,k+1} - I_{i,j,k}.$$

Then discretizing $(\star_1)$ and $(\star_2)$, we obtain

$$\frac{1}{h^2}\tilde{I}_{x,i,j,k}(\tilde{I}_{x,i,j,k}u_{i,j} + \tilde{I}_{y,i,j,k}v_{i,j} + \tilde{I}_{t,i,j,k}) - \frac{1}{h^2}\lambda(\bar{u}_{i,j} - 4u_{i,j}) = 0,$$
$$\frac{1}{h^2}\tilde{I}_{y,i,j,k}(\tilde{I}_{x,i,j,k}u_{i,j} + \tilde{I}_{y,i,j,k}v_{i,j} + \tilde{I}_{t,i,j,k}) - \frac{1}{h^2}\lambda(\bar{v}_{i,j} - 4v_{i,j}) = 0,$$

or equivalently,

$$(\tilde{I}^2_{x,i,j,k} + 4\lambda)u_{i,j} + (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})v_{i,j} = \lambda\bar{u}_{i,j} - \tilde{I}_{x,i,j,k}\tilde{I}_{t,i,j,k},$$
$$(\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})u_{i,j} + (\tilde{I}^2_{y,i,j,k} + 4\lambda)v_{i,j} = \lambda\bar{v}_{i,j} - \tilde{I}_{y,i,j,k}\tilde{I}_{t,i,j,k}.$$

Finite difference approximations (cont'd)

The system of equations can be rewritten as follows:

$$
\begin{aligned}
a u_{i,j} + b v_{i,j} &= r_1, \\
b u_{i,j} + c v_{i,j} &= r_2.
\end{aligned}
$$

By Cramer's rule, it yields

$$
u_{i,j} = \frac{\sigma_1}{\sigma}, \quad v_{i,j} = \frac{\sigma_2}{\sigma},
$$

where

$$
\begin{aligned}
\sigma = \begin{vmatrix} a & b \\ b & c \end{vmatrix} &= ac - b^2 = (\tilde{I}^2_{x,i,j,k} + 4\lambda)(\tilde{I}^2_{y,i,j,k} + 4\lambda) - (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})^2 \\
&= 4\lambda(\tilde{I}^2_{x,i,j,k} + \tilde{I}^2_{y,i,j,k} + 4\lambda),
\end{aligned}
$$

Finite difference approximations (cont'd)

and, moreover,

$$
\begin{aligned}
\sigma_1 = \begin{vmatrix} r_1 & b \\ r_2 & c \end{vmatrix} &= r_1 c - b r_2, \\
&= (\lambda \bar{u}_{i,j} - \tilde{I}_{x,i,j,k}\tilde{I}_{t,i,j,k})(\tilde{I}^2_{y,i,j,k} + 4\lambda) \\
&\quad - (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})(\lambda \bar{v}_{i,j} - \tilde{I}_{y,i,j,k}\tilde{I}_{t,i,j,k}) \\
&= \lambda \big((\tilde{I}^2_{y,i,j,k} + 4\lambda)\bar{u}_{i,j} - (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})\bar{v}_{i,j} - 4\tilde{I}_{x,i,j,k}\tilde{I}_{t,i,j,k} \big),
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma_2 = \begin{vmatrix} a & r_1 \\ b & r_2 \end{vmatrix} &= a r_2 - r_1 b, \\
&= (\tilde{I}^2_{x,i,j,k} + 4\lambda)(\lambda \bar{v}_{i,j} - \tilde{I}_{y,i,j,k}\tilde{I}_{t,i,j,k}) \\
&\quad - (\lambda \bar{u}_{i,j} - \tilde{I}_{x,i,j,k}\tilde{I}_{t,i,j,k})(\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k}) \\
&= \lambda \big((\tilde{I}^2_{x,i,j,k} + 4\lambda)\bar{v}_{i,j} - (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})\bar{u}_{i,j} - 4\tilde{I}_{y,i,j,k}\tilde{I}_{t,i,j,k} \big).
\end{aligned}
$$

The Jacobi iterative method

Then $u_{i,j}$ and $v_{i,j}$ are given by

$$u_{i,j} = \frac{(\tilde{I}^2_{y,i,j,k} + 4\lambda)\bar{u}_{i,j} - (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})\bar{v}_{i,j} - 4\tilde{I}_{x,i,j,k}\tilde{I}_{t,i,j,k}}{4(\tilde{I}^2_{x,i,j,k} + \tilde{I}^2_{y,i,j,k} + 4\lambda)}$$

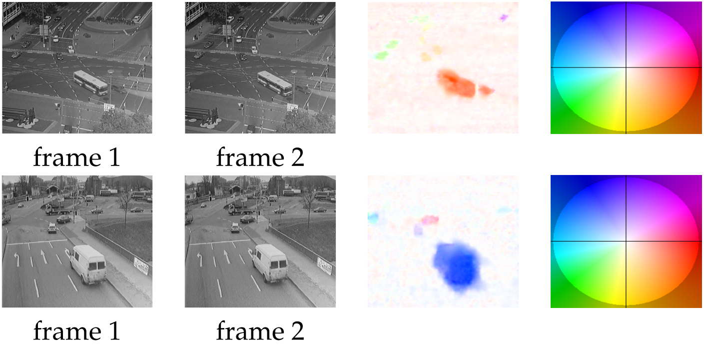
$$v_{i,j} = \frac{(\tilde{I}^2_{x,i,j,k} + 4\lambda)\bar{v}_{i,j} - (\tilde{I}_{x,i,j,k}\tilde{I}_{y,i,j,k})\bar{u}_{i,j} - 4\tilde{I}_{y,i,j,k}\tilde{I}_{t,i,j,k}}{4(\tilde{I}^2_{x,i,j,k} + \tilde{I}^2_{y,i,j,k} + 4\lambda)}$$

We can solved it by the Jacobi iterative method:

$$u^{(\ell+1)}_{i,j} = \frac{\bar{u}^{(\ell)}_{i,j}}{4} - \frac{\tilde{I}_{x,i,j,k}\left(\tilde{I}_{x,i,j,k}\bar{u}^{(\ell)}_{i,j} + \tilde{I}_{y,i,j,k}\bar{v}^{(\ell)}_{i,j} + 4\tilde{I}_{t,i,j,k}\right)}{4\left(\tilde{I}^2_{x,i,j,k} + \tilde{I}^2_{y,i,j,k} + 4\lambda\right)},$$

$$v^{(\ell+1)}_{i,j} = \frac{\bar{v}^{(\ell)}_{i,j}}{4} - \frac{\tilde{I}_{y,i,j,k}\left(\tilde{I}_{y,i,j,k}\bar{v}^{(\ell)}_{i,j} + \tilde{I}_{x,i,j,k}\bar{u}^{(\ell)}_{i,j} + 4\tilde{I}_{t,i,j,k}\right)}{4\left(\tilde{I}^2_{x,i,j,k} + \tilde{I}^2_{y,i,j,k} + 4\lambda\right)}.$$

Experimental results of the Horn-Schunck model



References:

- **[HS-1981]** B. K. P. Horn, and B. G. Schunck, Determining optical flow, *Artificial Intelligence,* 17 (1981), pp. 185-203.
- **[GN-1987]** M. A. Gennert, and S. Negahdaripour, Relaxing the brightness constancy assumption in computing optical flow, *A. I. Memo No. 975,* (1987).