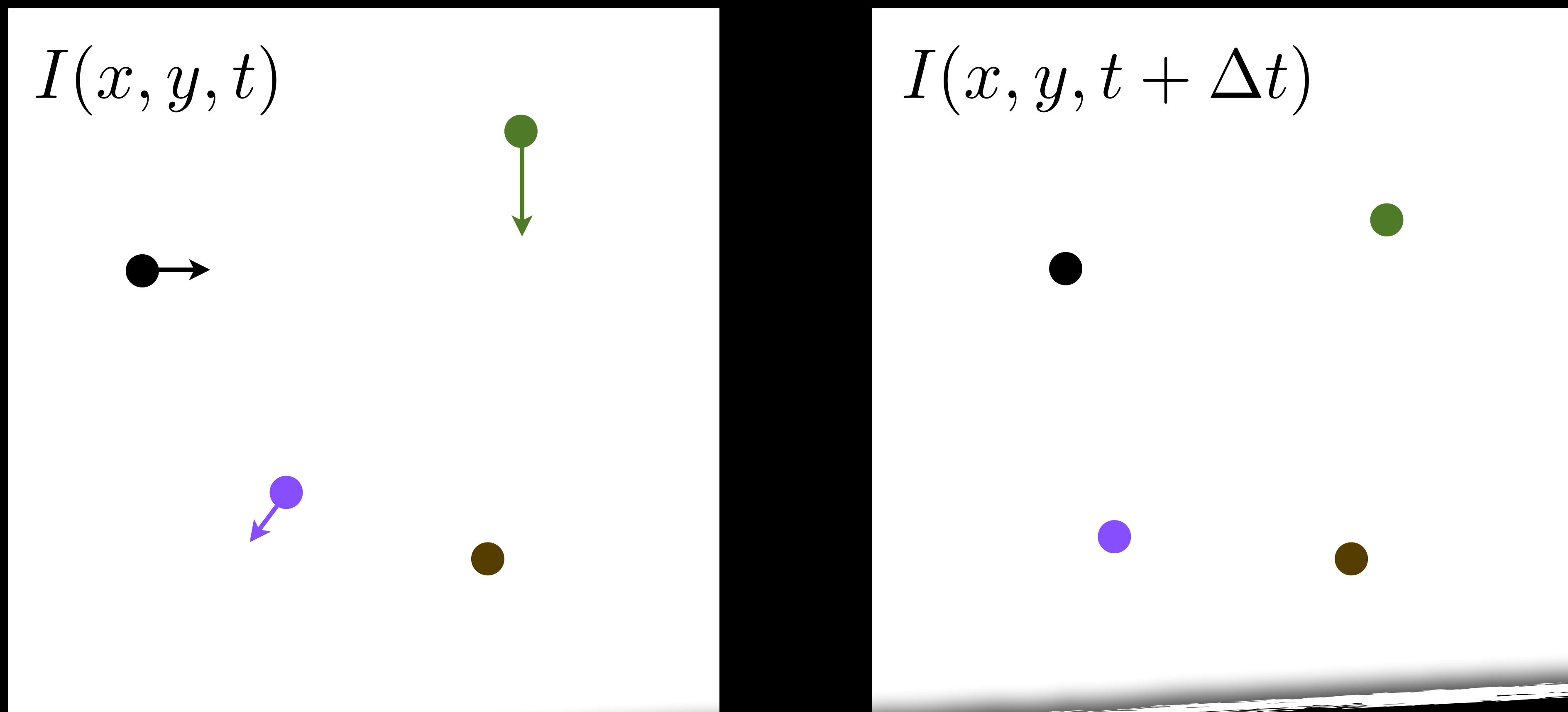
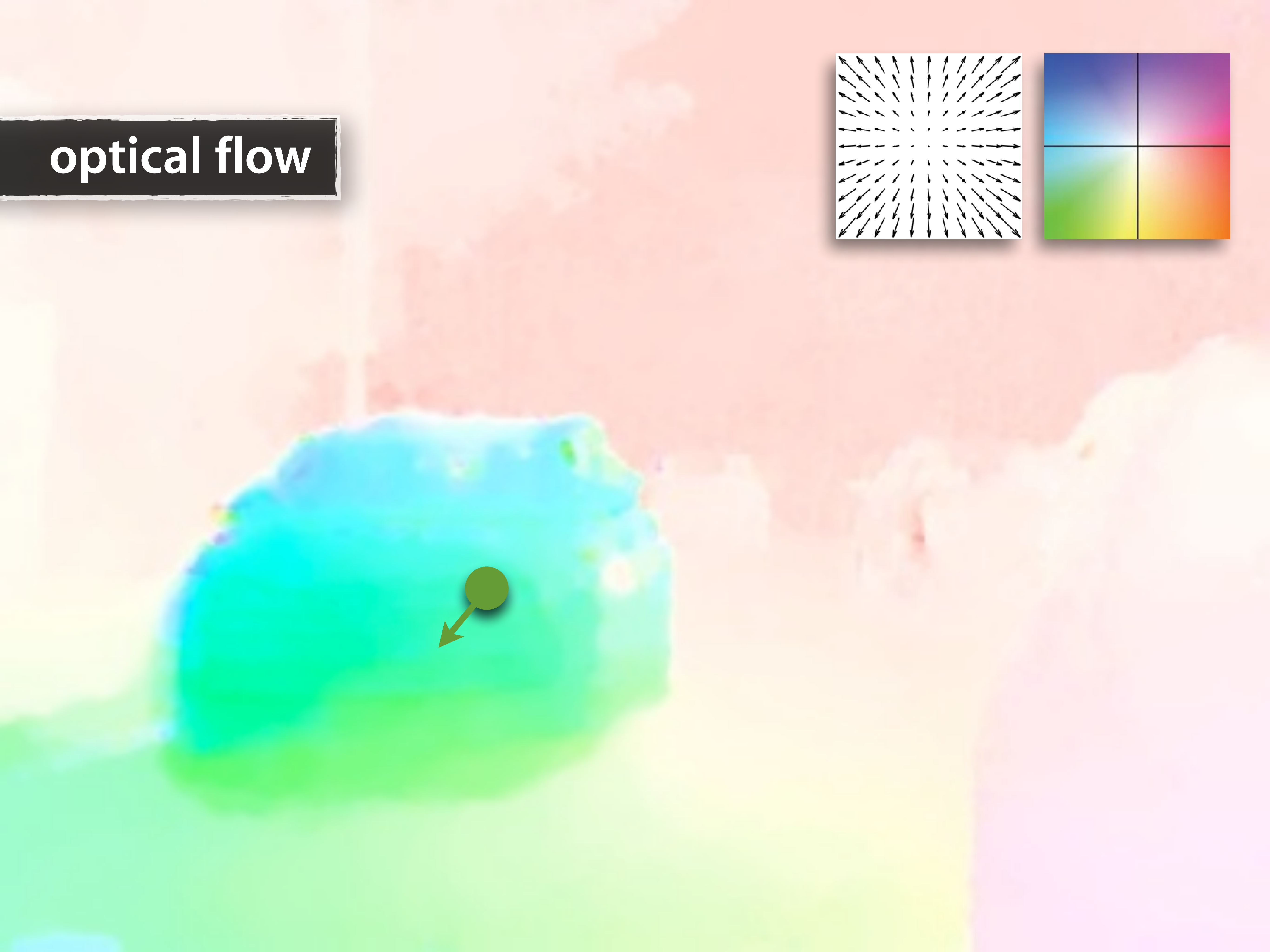
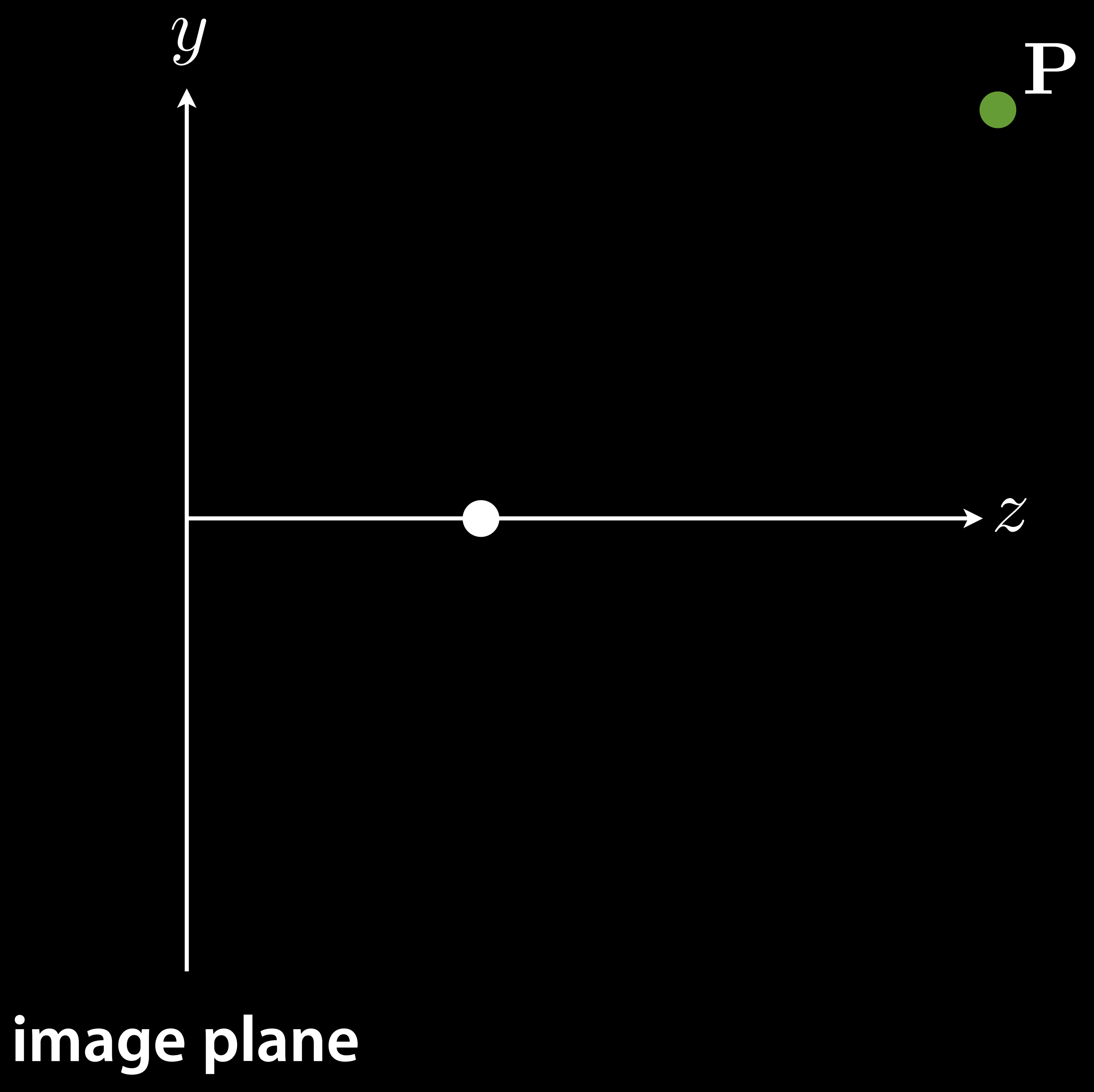
$f : \mathbb{R}^3 \to \mathbb{R}$ **or** $f : \mathbb{R}^3 \to \mathbb{R}^3$

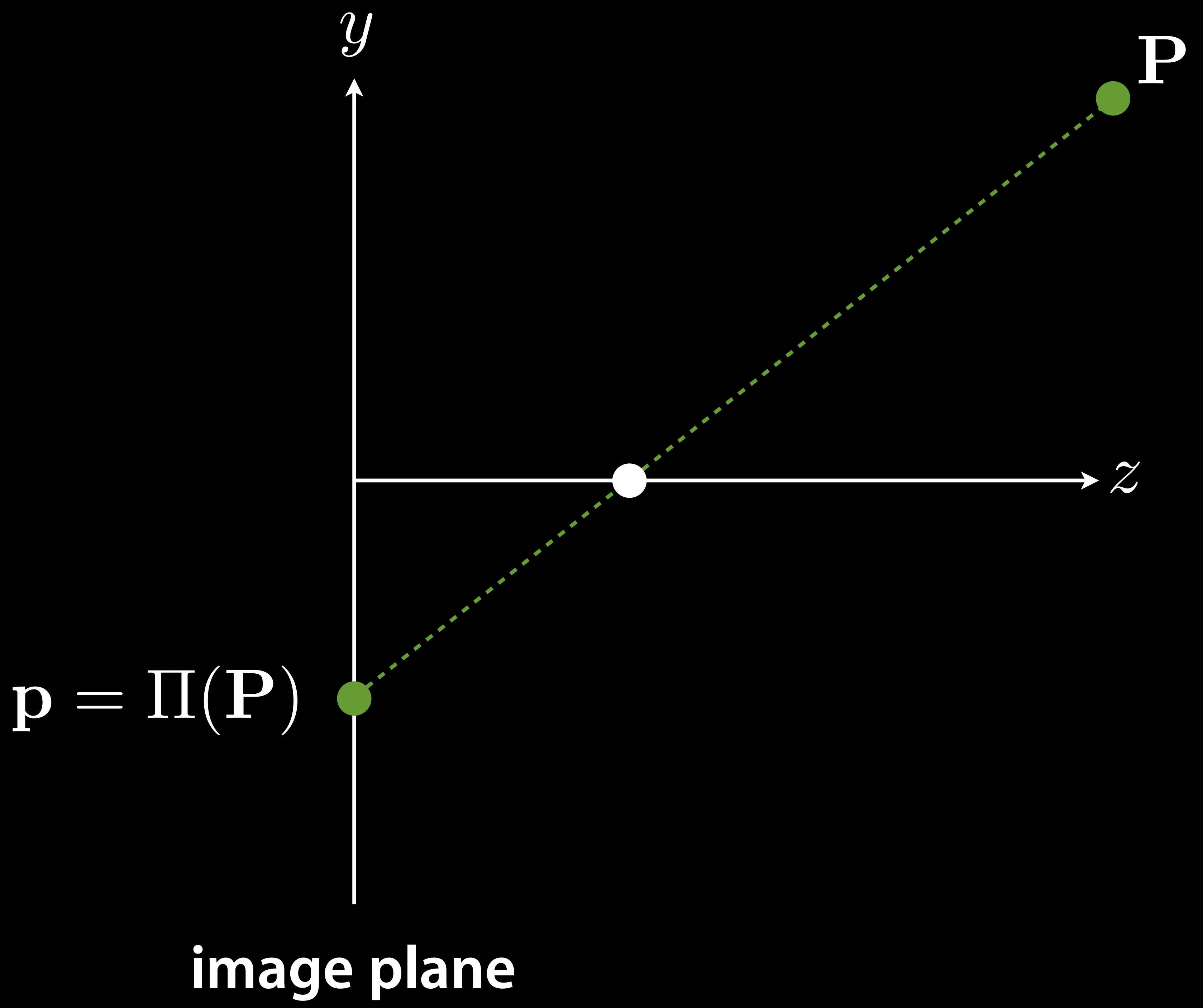
Problem Definition



GOAL: Estimate pixel displacement between images

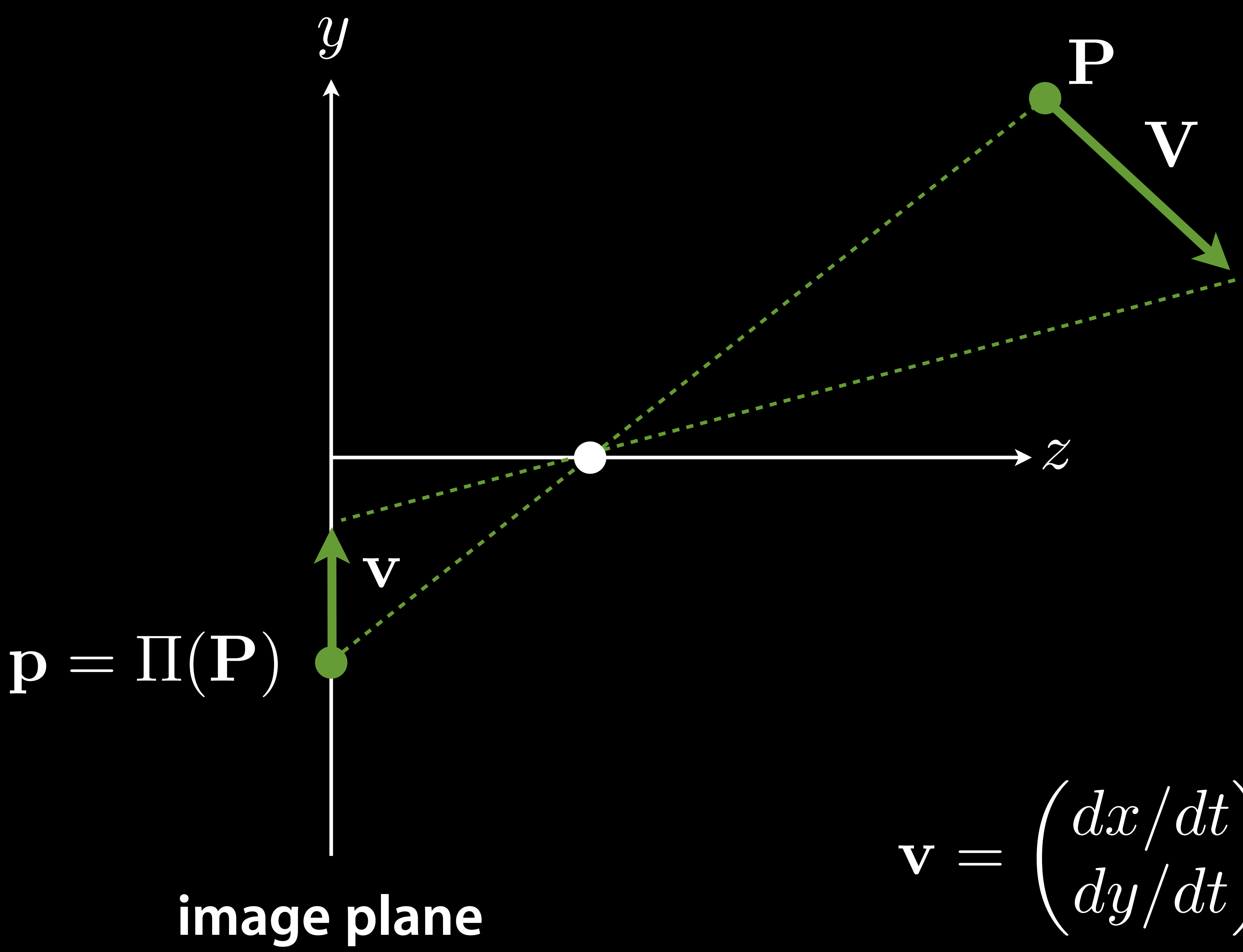

optical flow


y
$\mathbf{P}$
z
image plane


y
$\mathbf{P}$
z
$\mathbf{p} = \Pi(\mathbf{P})$
image plane

Motion Field

projection of the 3D scene velocities onto the image plane



$$\mathbf{v} = \begin{pmatrix} dx/dt \\ dy/dt \end{pmatrix} = \frac{d\mathbf{p}}{dt}$$

$$= \frac{d\Pi(\mathbf{P})}{dt}$$

Optical Flow

distribution of apparent velocities of movement of brightness patterns in the image

motion derived from image measurements

Motion Field

$$\neq$$

Optical Flow

Motion Field

$$\approx$$

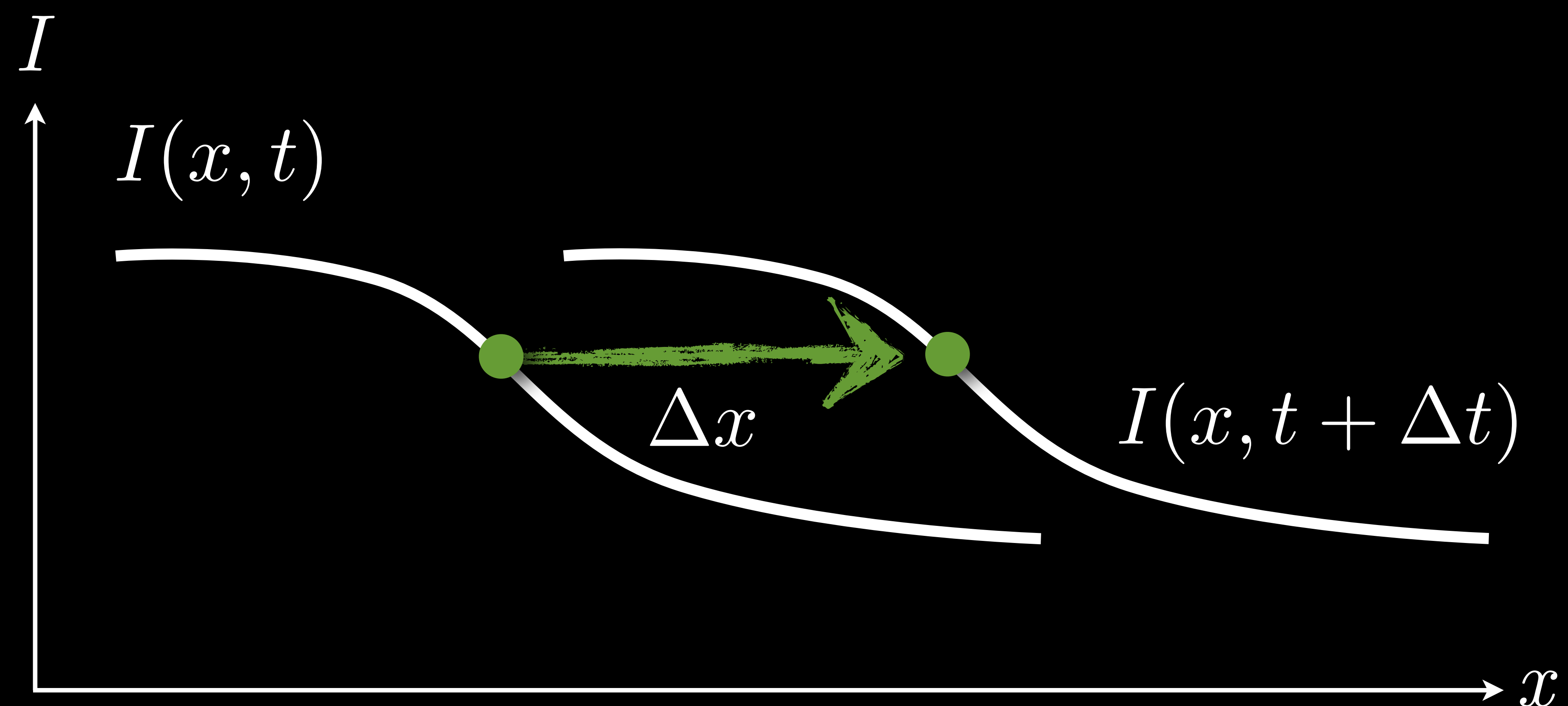
Optical Flow

Assumption

Image brightness of a point remains constant over time

other image properties are possible

Brightness Constancy



Brightness Constancy Assumption

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

Brightness Constancy Assumption

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

assuming small duration

Review: Taylor series

Taylor Series Expansion

$$
\begin{aligned}
f(x+\Delta x, y+\Delta y, z+\Delta z) &= f(x,y,z) \\
&\quad + f_x(x,y,z)\Delta x + f_y(x,y,z)\Delta y + f_z(x,y,z)\Delta z \\
&\quad + \text{higher order terms}
\end{aligned}
$$

Brightness Constancy Assumption

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

since image is non-linear in flow, flow recovery is difficult

Brightness Constancy Assumption

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

Taylor series

$$\cancel{I(x,y,t)} = \cancel{I(x,y,t)} + \frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \text{h.o.t.}$$

simplify

$$\frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \text{h.o.t.} = 0$$

divide by Δt

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} + \text{h.o.t.} = 0$$

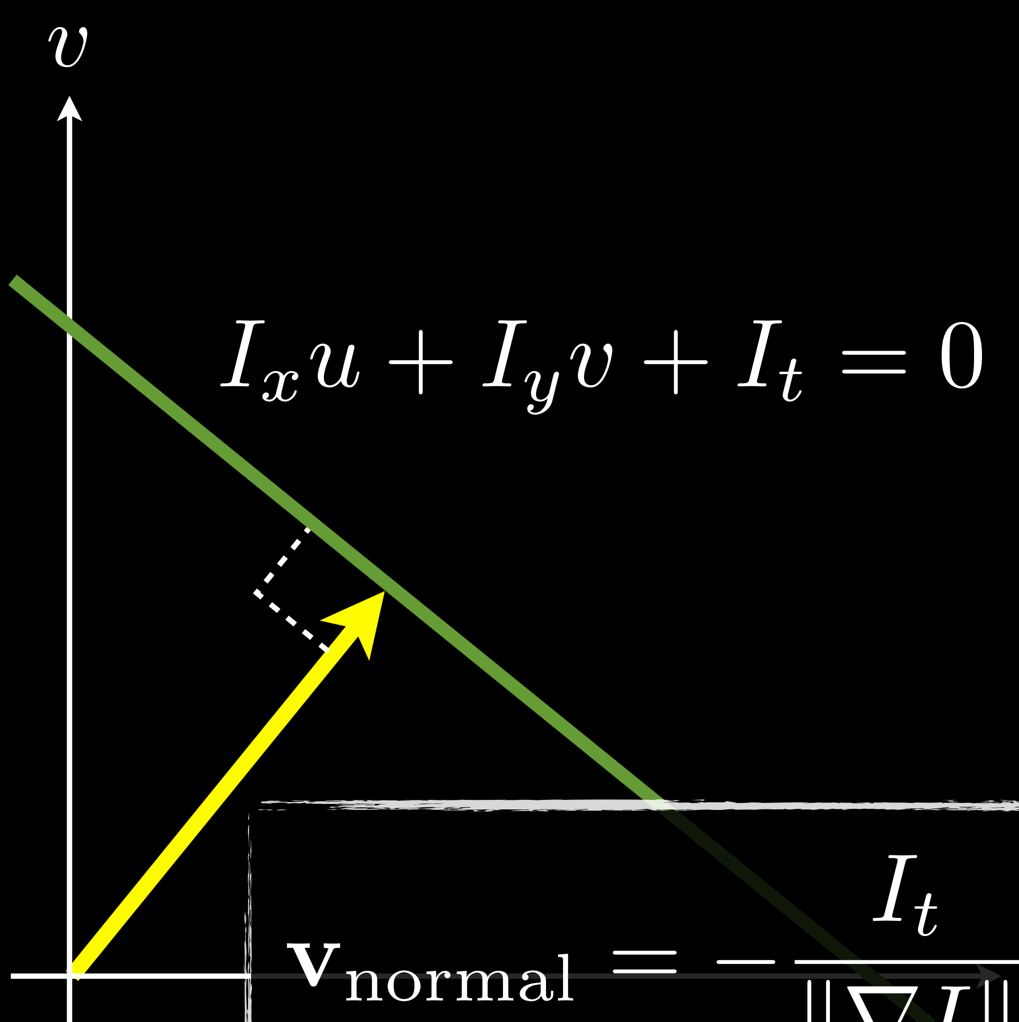
$\Delta t \rightarrow 0$

Brightness Constancy Constraint

$$I_x u + I_y v + I_t = 0$$

Brightness Constancy Constraint

$$I_x u + I_y v + I_t = 0$$

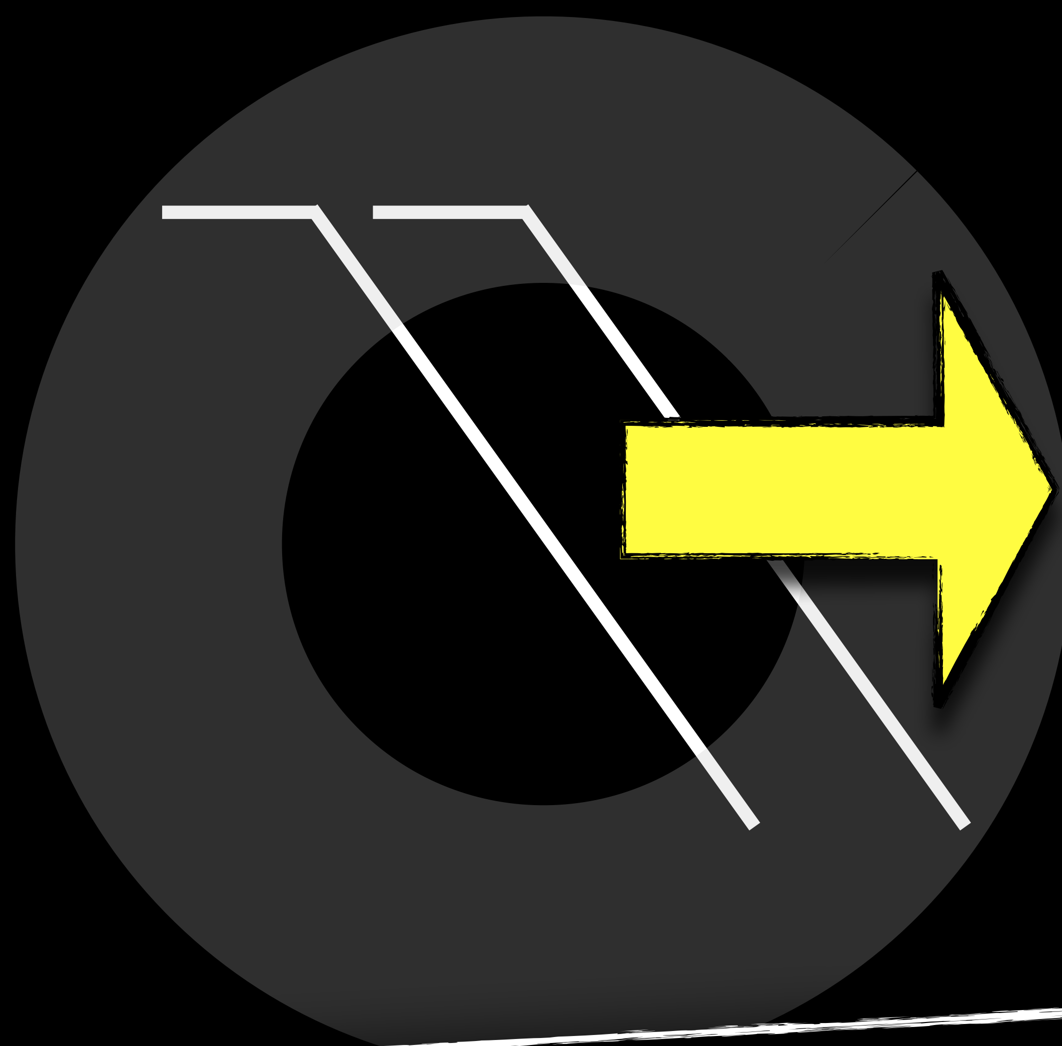


$$\mathbf{u}^\top \nabla I = -I_t$$

$$\mathbf{u}^\top \frac{\nabla I}{\|\nabla I\|} = -\frac{I_t}{\|\nabla I\|}$$

$$\mathbf{v}_{\text{normal}} = -\frac{I_t}{\|\nabla I\|} \frac{\nabla I}{\|\nabla I\|}$$

normal flow



Which way are the lines moving?

Estimation

Optical Flow


Local
Global

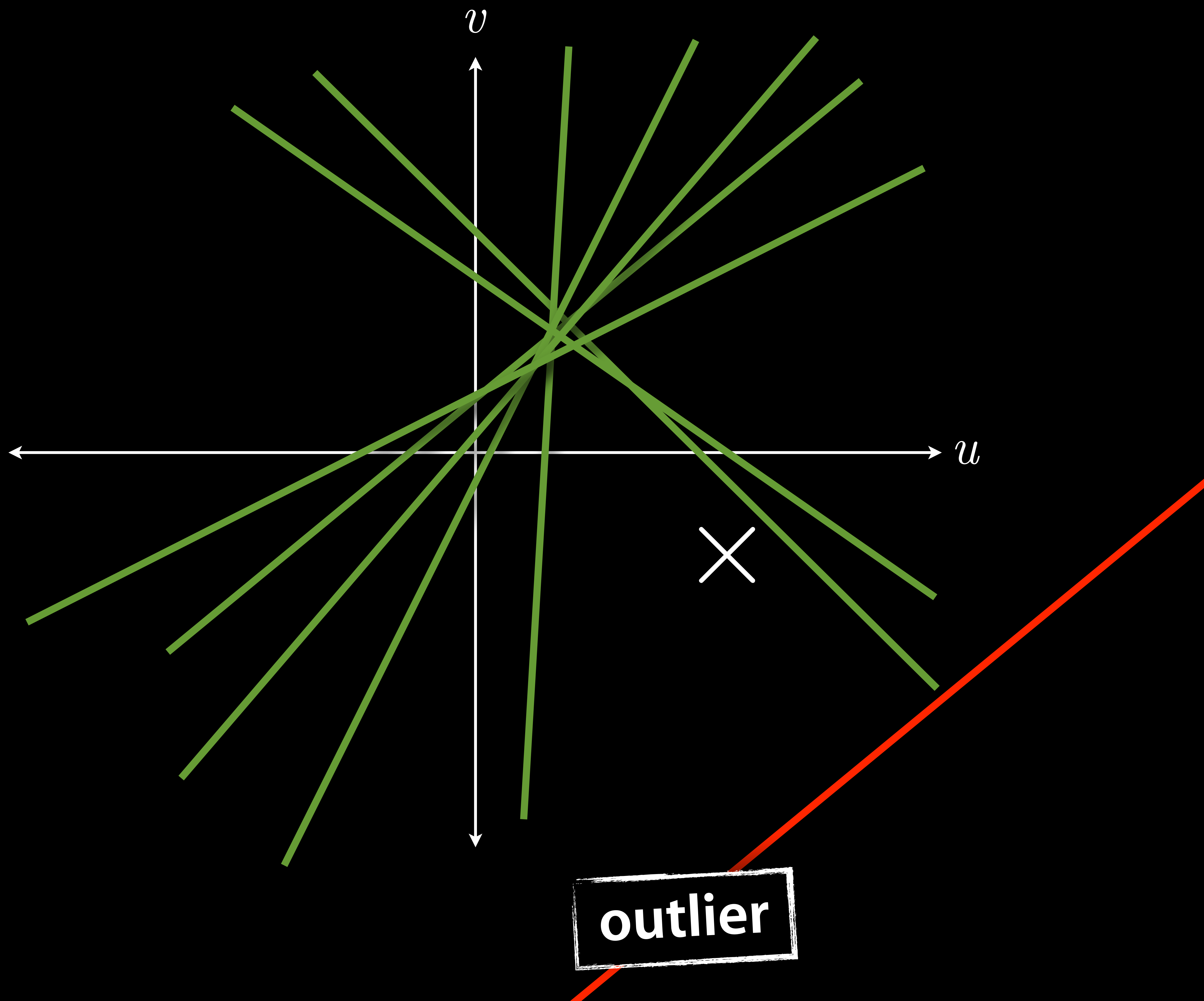

Assume constant velocity in a pixel's neighbourhood
$N \times N$

Lucas-Kanade

$$\underbrace{\begin{pmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{N^2}) & I_y(\mathbf{p}_{N^2}) \end{pmatrix}}_{\mathbf{A}_{N^2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = - \underbrace{\begin{pmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{N^2}) \end{pmatrix}}_{\mathbf{b}_{N^2\times 1}}$$

How can we solve this system?

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$



v
u
outlier

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^\top \mathbf{A}\mathbf{v} = \mathbf{A}^\top \mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^\top \mathbf{A}_{2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = - \underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^\top \mathbf{b}_{2\times 1}}$$

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^\top \mathbf{A}\mathbf{v} = \mathbf{A}^\top \mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^\top \mathbf{A}_{2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = - \underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^\top \mathbf{b}_{2\times 1}}$$

Does this look familiar?

Harris Corners

$$E(\Delta x, \Delta y) = \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} \mathbf{M} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

$$\mathbf{M} = \sum_{x,y} w(x,y) \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

M captures the structure within the local patch

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^\top \mathbf{A}\mathbf{v} = \mathbf{A}^\top \mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^\top\mathbf{A}_{2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = - \underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^\top\mathbf{b}_{2\times 1}}$$

$$\mathbf{v} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$$

$$\mathbf{v} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}$$

Conditions for solvability?

$\mathbf{A}^\top \mathbf{A}$ **is invertible**

$\mathbf{A}^\top \mathbf{A}$ **eigenvalues** $\lambda_1, \lambda_2 \gg 0$

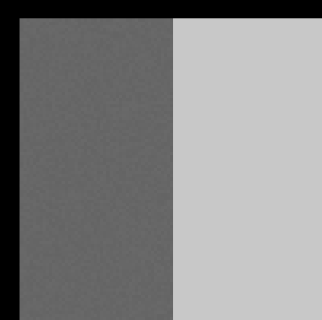
Conditions for solvability?

$\mathbf{A}^\top \mathbf{A}$ **is invertible**

$\mathbf{A}^\top \mathbf{A}$ **eigenvalues** $\lambda_1, \lambda_2 \gg 0$

$\mathbf{A}^\top \mathbf{A}$ **should be well conditioned,** λ_1 / λ_2 **not too large**

“Flat”



“Edge”

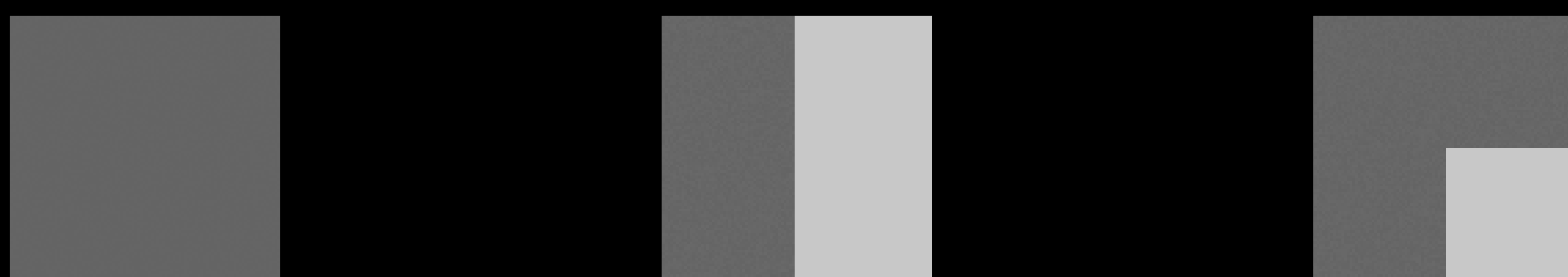
“Corner”

Conditions for solvability?

$\mathbf{A}^\top\mathbf{A}$ is invertible

$\mathbf{A}^\top\mathbf{A}$ eigenvalues $\lambda_1, \lambda_2 \gg 0$

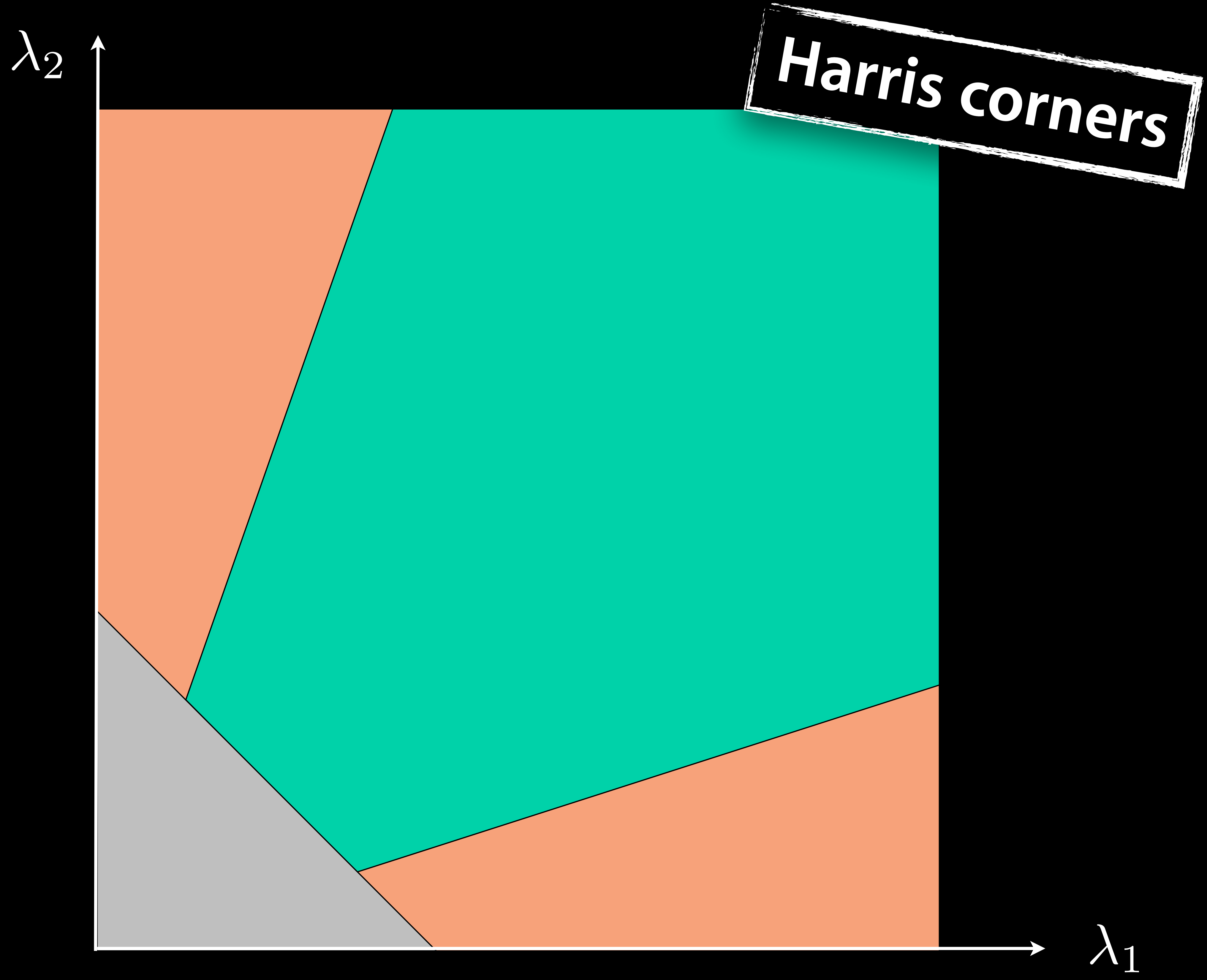
$\mathbf{A}^\top\mathbf{A}$ should be well conditioned, λ_1/λ_2 not too large



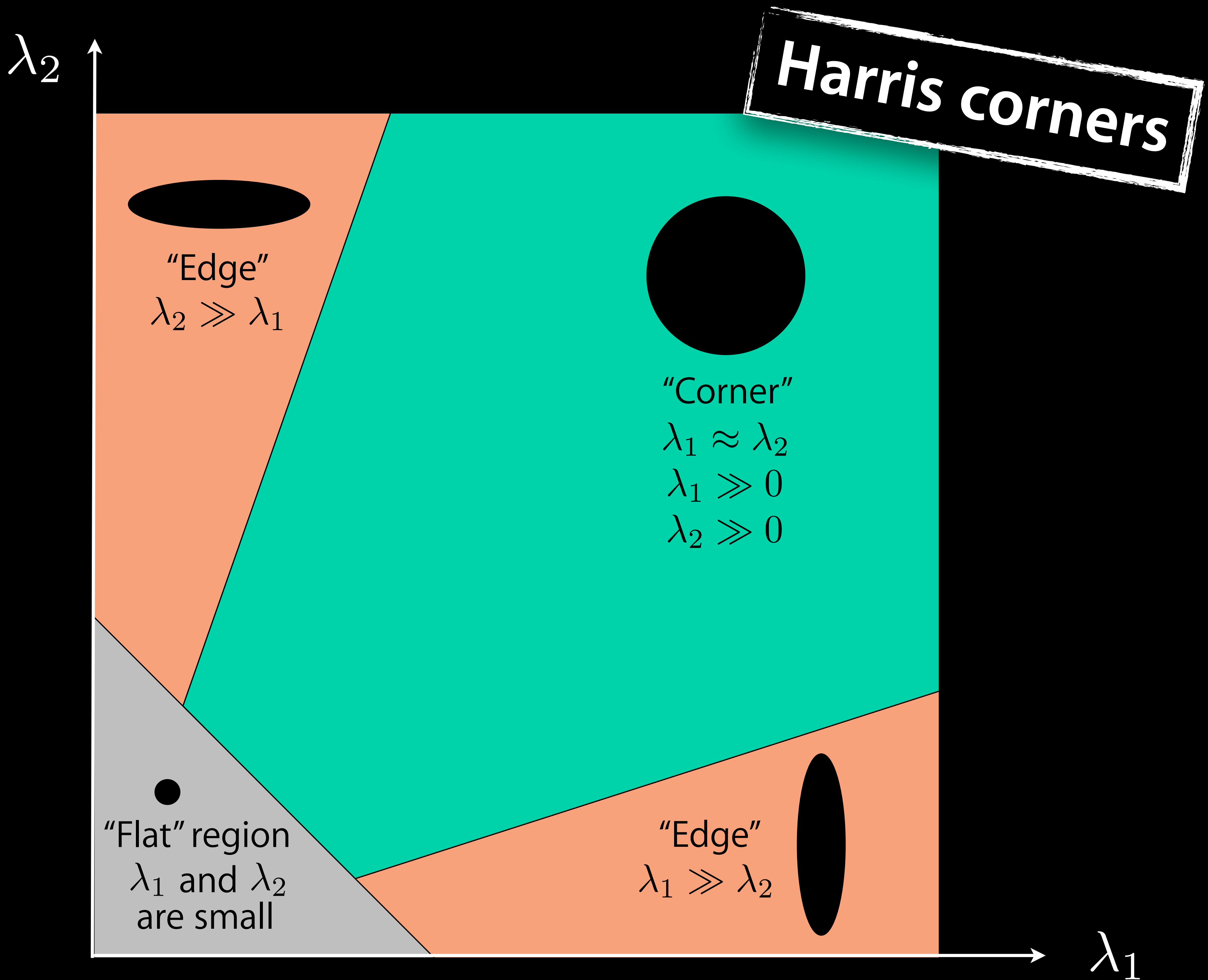
"Flat" "Edge" "Corner"

Which pattern satisfies these conditions?

Classification of image points using eigenvalues of M



Classification of image points using eigenvalues of M

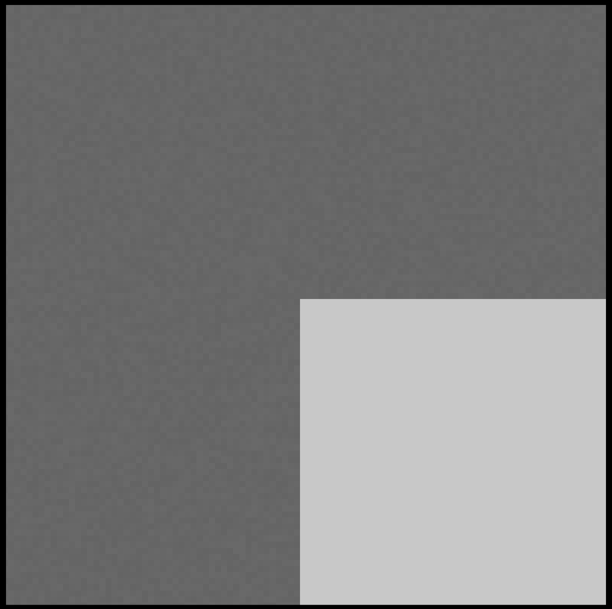


Conditions for solvability?

$\mathbf{A}^\top \mathbf{A}$ is invertible

$\mathbf{A}^\top \mathbf{A}$ eigenvalues $\lambda_1, \lambda_2 \gg 0$

$\mathbf{A}^\top \mathbf{A}$ should be well conditioned, λ_1 / λ_2 not too large



“Corner”

Pros

Easy to implement

Fast computations

Cons

- Fails when brightness is not constant
- Fails when intensity structure within window is poor
- Fails along boundaries
- Fails when the velocity is large
- Locally constant translation is restrictive

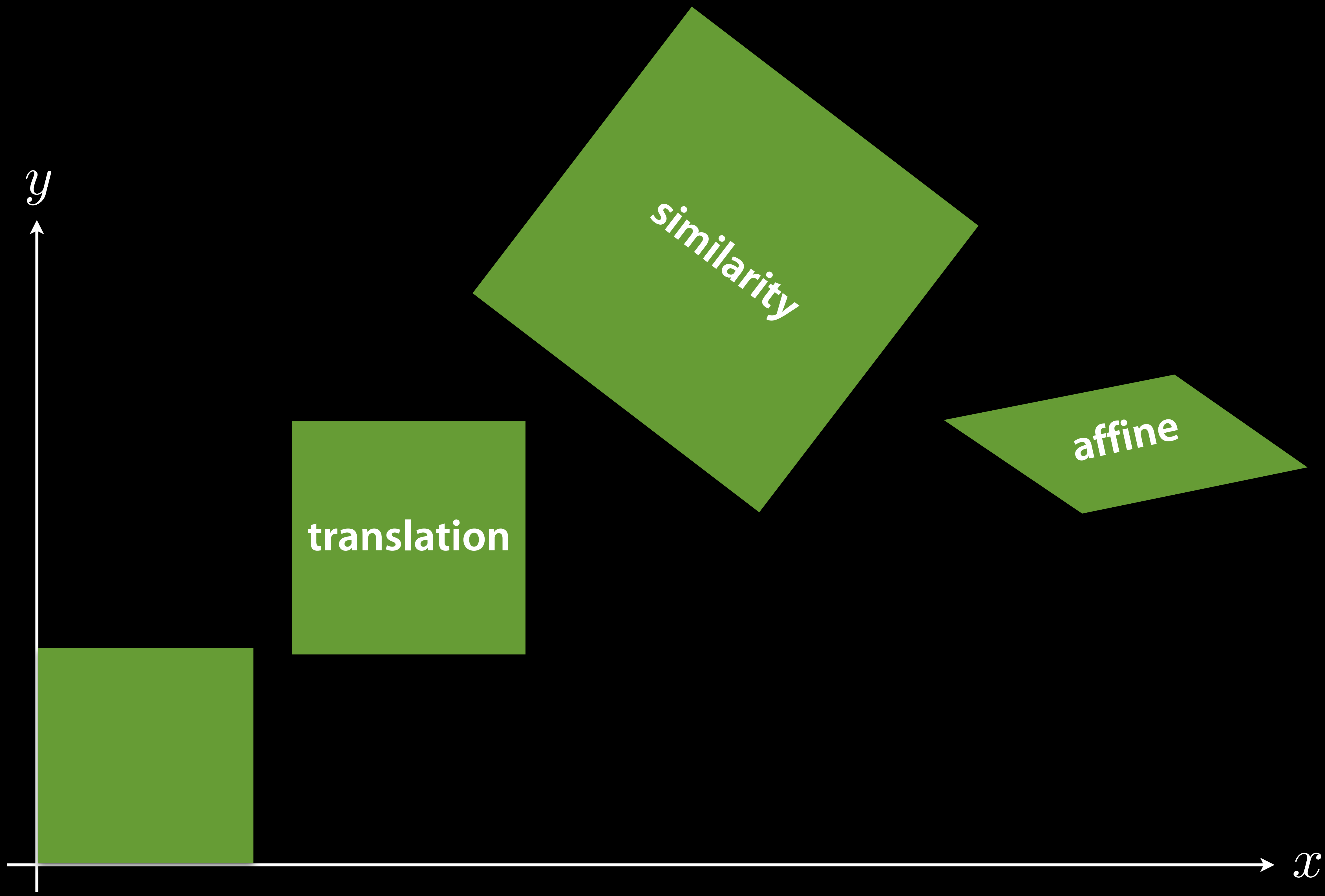
$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$$

translational patch motion is restrictive

$$I(x, y, t) = I(x + \Delta x(\mathbf{x}; \mathbf{p}), y + \Delta y(\mathbf{x}; \mathbf{p}), t + \Delta t)$$

motion described by a parametric model

Parametric motion examples



Affine Motion

$$u(x, y) = a_1 + a_2 x + a_3 y$$
$$v(x, y) = a_4 + a_5 x + a_6 y$$

substitute into brightness constancy equation

$$I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \approx 0$$

one linear constraint in six unknowns per pixel

solve over pixel neighbourhood using least-squares

Lucas-Kanade Optical Flow [Lucas, Kanade 81]

minimize $E(\boldsymbol{u}) = \sum_{x \in W} (G(\boldsymbol{x}) - F(\boldsymbol{x}+\boldsymbol{u}))^2$

Small $\boldsymbol{u}$ → **Taylor-expand** $F(x+u)$

$$F(\boldsymbol{x}+\boldsymbol{u}) \cong F(\boldsymbol{x}) + \frac{\partial F}{\partial \boldsymbol{x}}^T \boldsymbol{u}$$

→ $E(\boldsymbol{u})$ becomes quadratic of $\boldsymbol{u}$

→ $\boldsymbol{u}$ can be obtained explicitly as

$$\boldsymbol{u} = \sum (G(\boldsymbol{x}) - F(\boldsymbol{x})) \frac{\partial F}{\partial \boldsymbol{x}} \left[\sum \frac{\partial F}{\partial \boldsymbol{x}} \frac{\partial F}{\partial \boldsymbol{x}}^T \right]^{-1}$$

KLT Tracking Summary

1. **Find good features to track**

    Harris corners or threshold on smallest eigenvalue

2. **Compute frame-to-frame feature displacement**

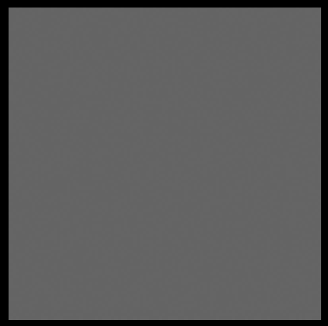
    estimate patch translation using Lucas-Kanade

3. **Verify appearance of patch**

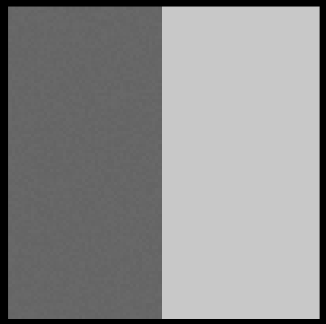
    compute affine registration between current patch and patch in first frame


Global

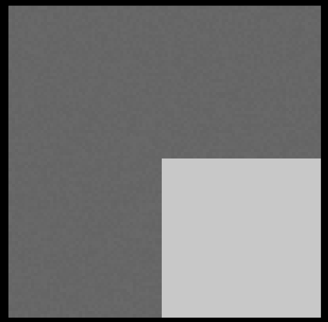
Local motion is inherently ambiguous



totally ambiguous



definite along the normal,
ambiguous along the tangent

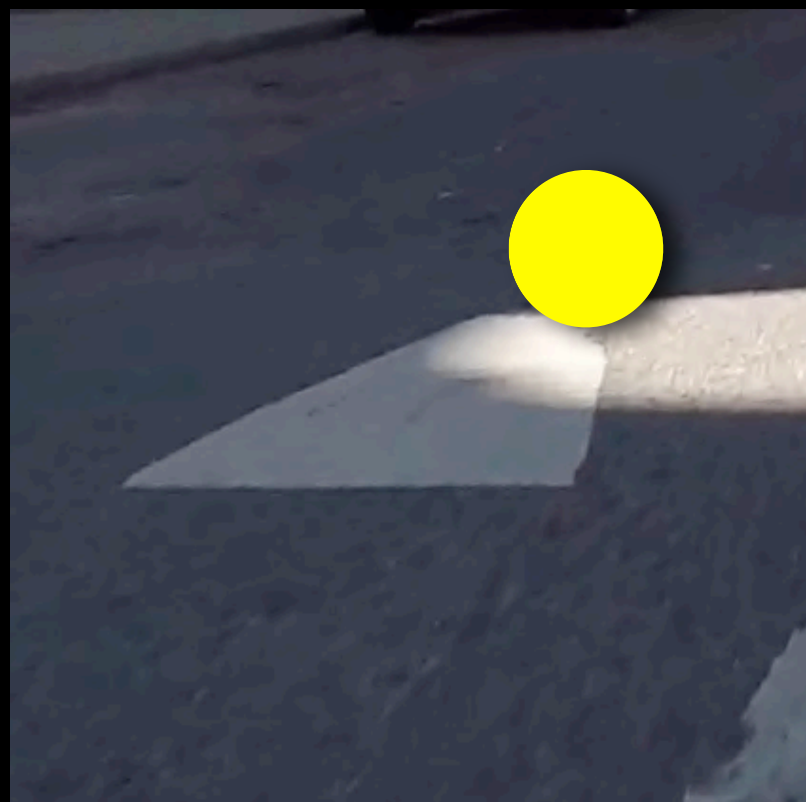


no ambiguity

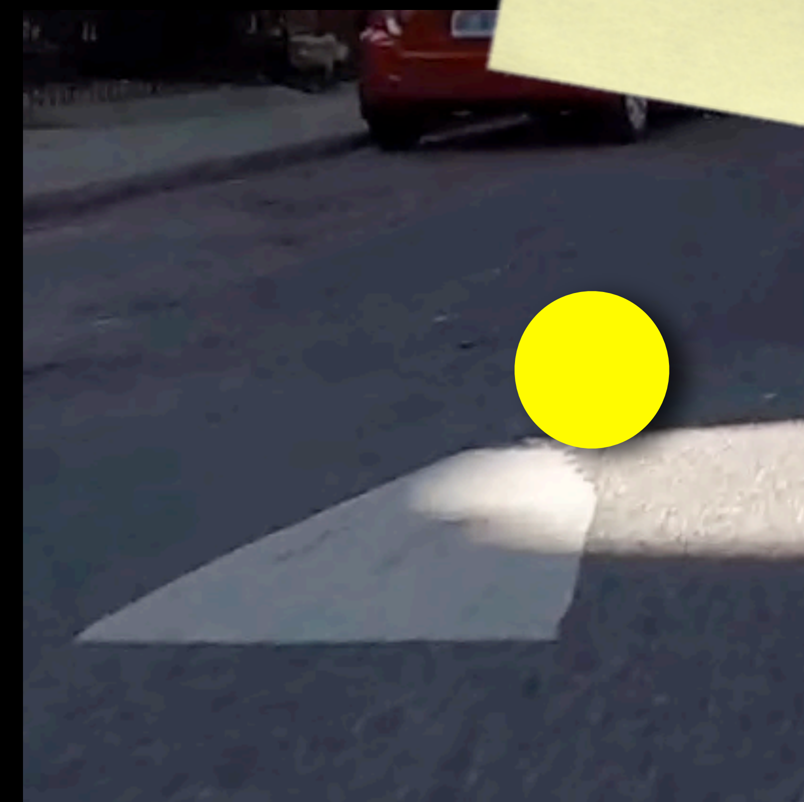
2

assumptions

Assumption 1



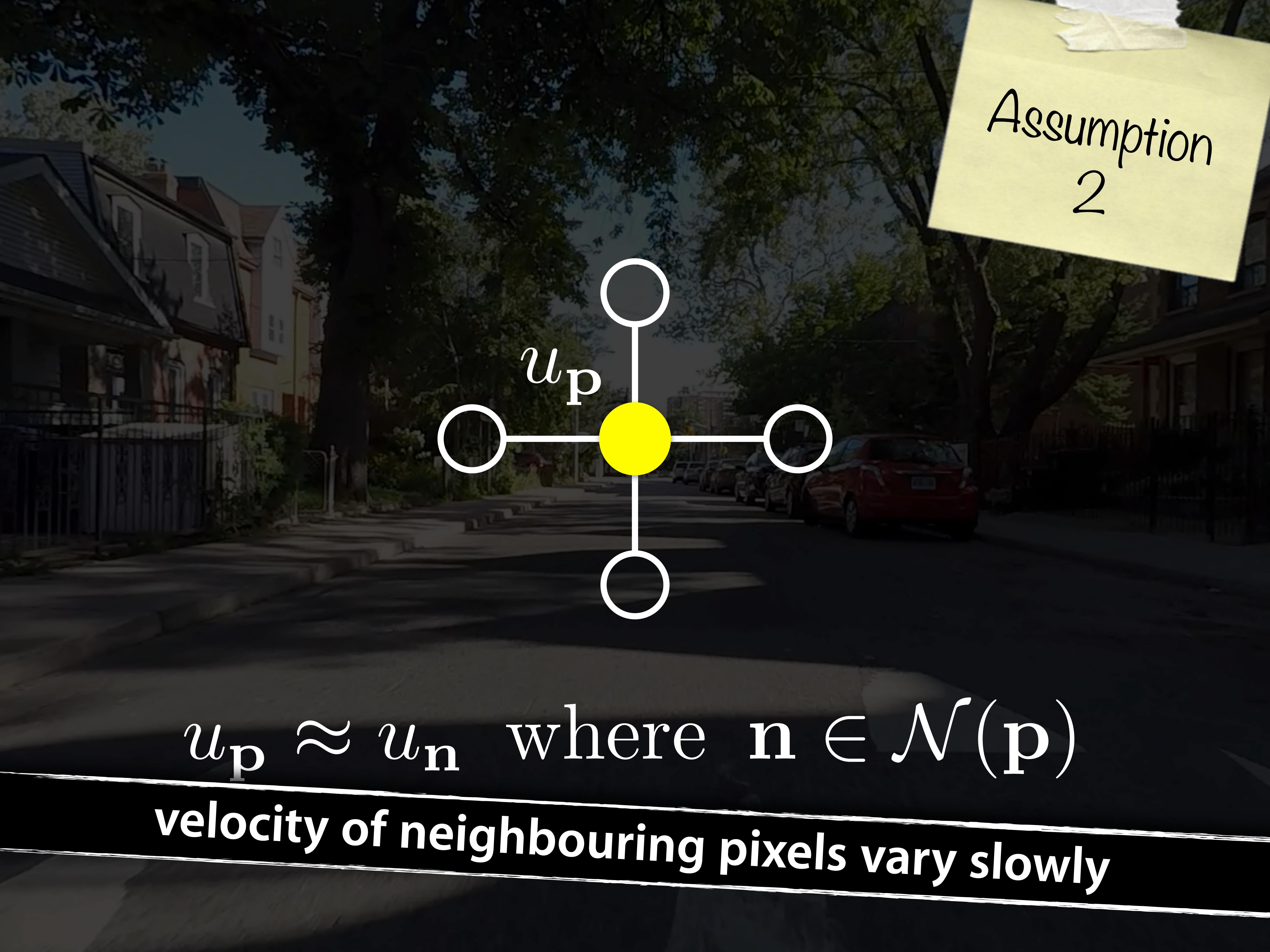
frame t



frame $t+1$

$$I(x, y, t) = I(x+u, y+v, t+1)$$

appearance constancy assumption


Assumption 2
$u_{\mathbf{p}}$
$u_{\mathbf{p}} \approx u_{\mathbf{n}}$ where $\mathbf{n} \in \mathcal{N}(\mathbf{p})$
velocity of neighbouring pixels vary slowly

Photometric Loss + Smoothness Loss

minimize loss with respect to velocity field over image pair

Horn-Schunck

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \{(I_x u_{i,j} + I_y v_{i,j} + I_t)^2$$
$$+\lambda[(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2$$
$$+(v_{i,j} - v_{i+1,j})^2 + (v_{i,j} - v_{i,j+1})^2]\}$$

each pixel is associated with its own velocity variable

data term

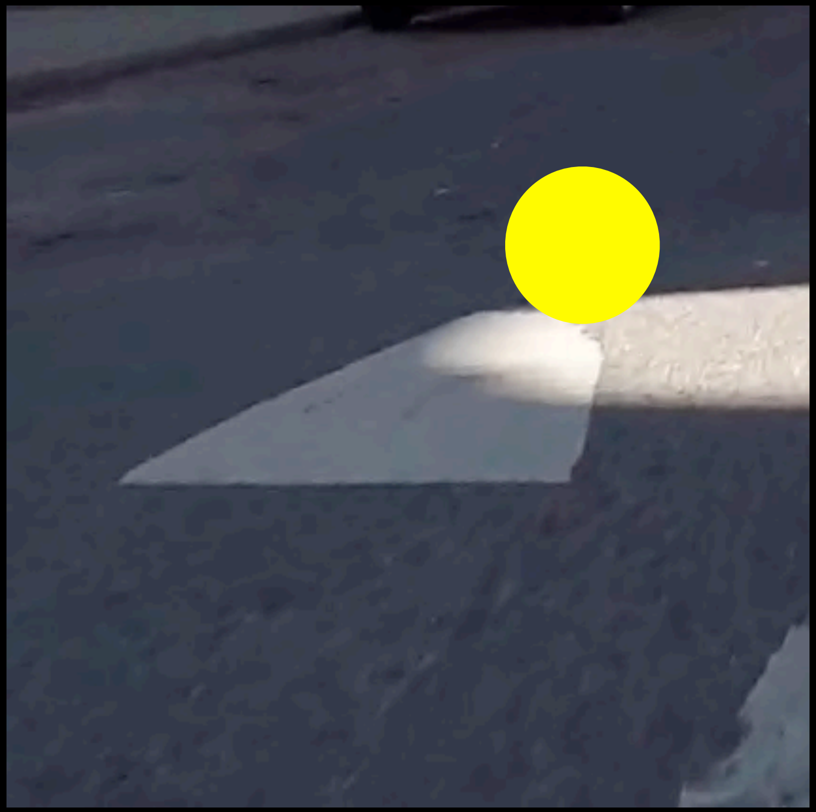
$$\sum\{(I_x u_{i,j} + I_y v_{i,j} + I_t)^2$$

measures departure from brightness constancy

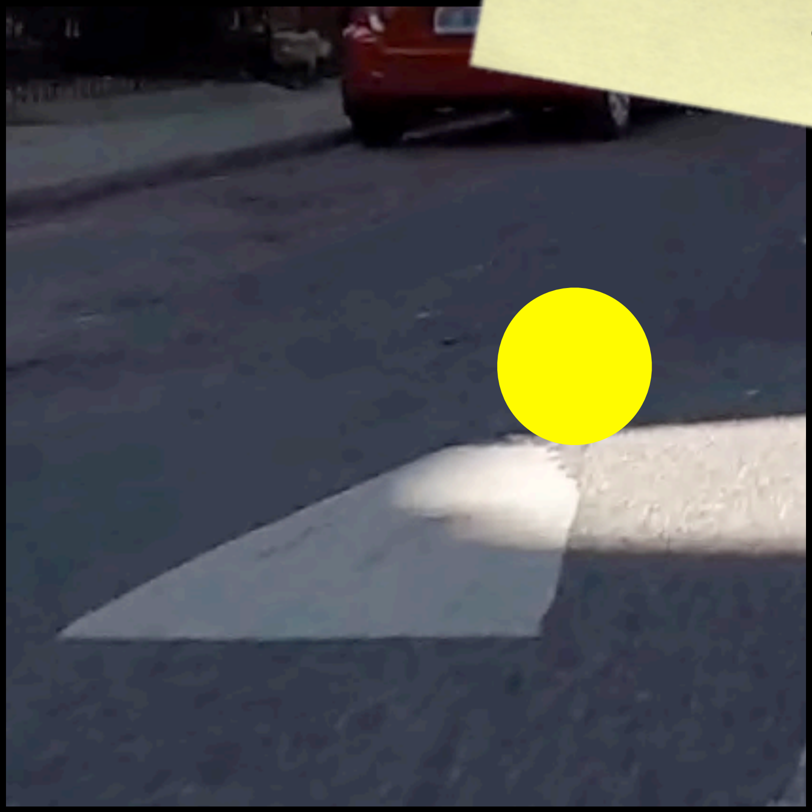
$$+\lambda[(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} -$$

$$+(v_{i,j} - v_{i+1,j})^2 + (v_{i,j} - v$$

Assumption 1



frame t



frame $t+1$

$$I(x, y, t) = I(x+u, y+v, t+1)$$

appearance constancy assumption

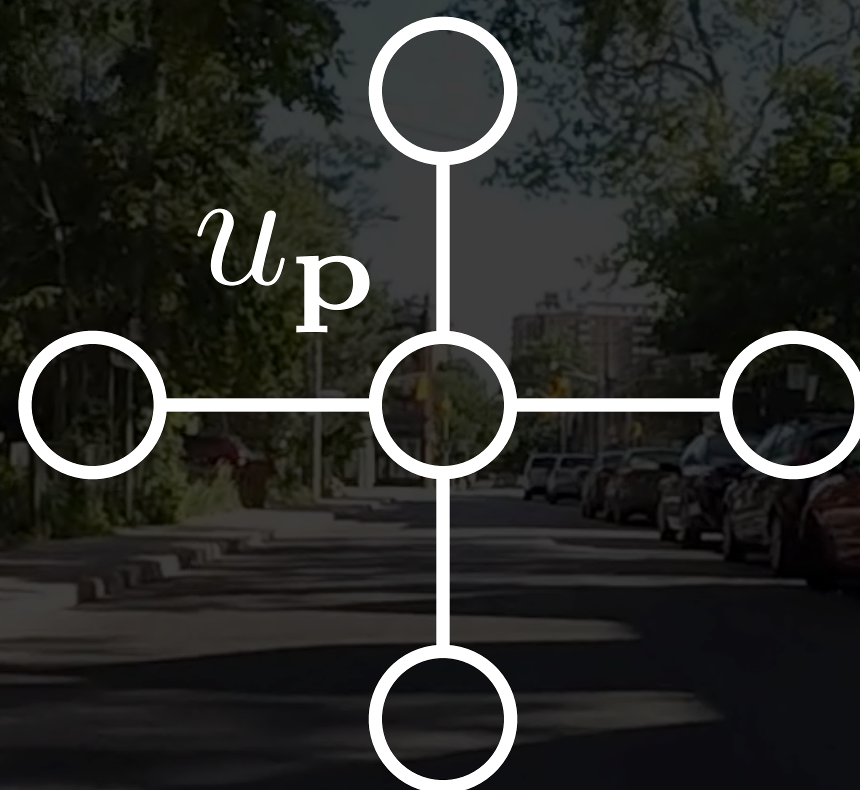
$$n \sum_{i,j} \{(I_x u_{i,j} + \ldots)^2$$

smoothness term

$$+\lambda[(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2$$

$$+(v_{i,j} - v_{i+1,j})^2 + (v_{i,j} - v_{i,j+1})^2]\}$$

Assumption 2



$$u_{\mathbf{p}} \approx u_{\mathbf{n}} \quad \text{where} \quad \mathbf{n} \in \mathcal{N}(\mathbf{p})$$

velocity of neighbouring pixels vary slowly

$$n \sum_{i,j} \{ (I_x u_{i,j} \ldots)^2$$

smoothness term

$$+\lambda[(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2$$

$$+(v_{i,j} - v_{i+1,j})^2 + (v_{i,j} - v_{i,j+1})^2]\}$$

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \{(I_x u_{i,j} + I_y v_{i,j} + I_t)^2$$
$$+ \lambda [(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2$$
$$+ (v_{i,j} - v_{i+1,j})^2 + (v_{i,j} - v_{i,j+1})^2]\}$$

relative weighting


data term
smoothness term

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \{(I_x u_{i,j} + I_y v_{i,j} + I_t)^2$$
$$+\lambda[(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2$$
$$+(v_{i,j} - v_{i+1,j})^2 + (v_{i,j} - v_{i,j+1})^2]\}$$

Differentiate by unknowns and set to zero

$$(I_x u_{i,j} + I_y v_{i,j} + I_t)I_x +$$
$$\lambda[4u_{i,j} - (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1})] = 0$$

$$(I_x u_{i,j} + I_y v_{i,j} + I_t)I_y +$$
$$\lambda[4v_{i,j} - (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1})] = 0$$

$$(I_x u_{i,j} + I_y v_{i,j} + I_t) I_x +$$
$$\lambda[4u_{i,j} - (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1})] = 0$$

$$(I_x u_{i,j} + I_y v_{i,j} + I_t) I_y +$$
$$\lambda[4v_{i,j} - (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1})] = 0$$

each image point yields a pair of linear equations

$$(I_x u_{i,j} + I_y v_{i,j} + I_t)I_x + \lambda[4u_{i,j} - (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1})] = 0$$

$$(I_x u_{i,j} + I_y v_{i,j} + I_t)I_y + \lambda[4v_{i,j} - (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1})] = 0$$

equations can be solved using Gaussian elimination

in theory

Pros

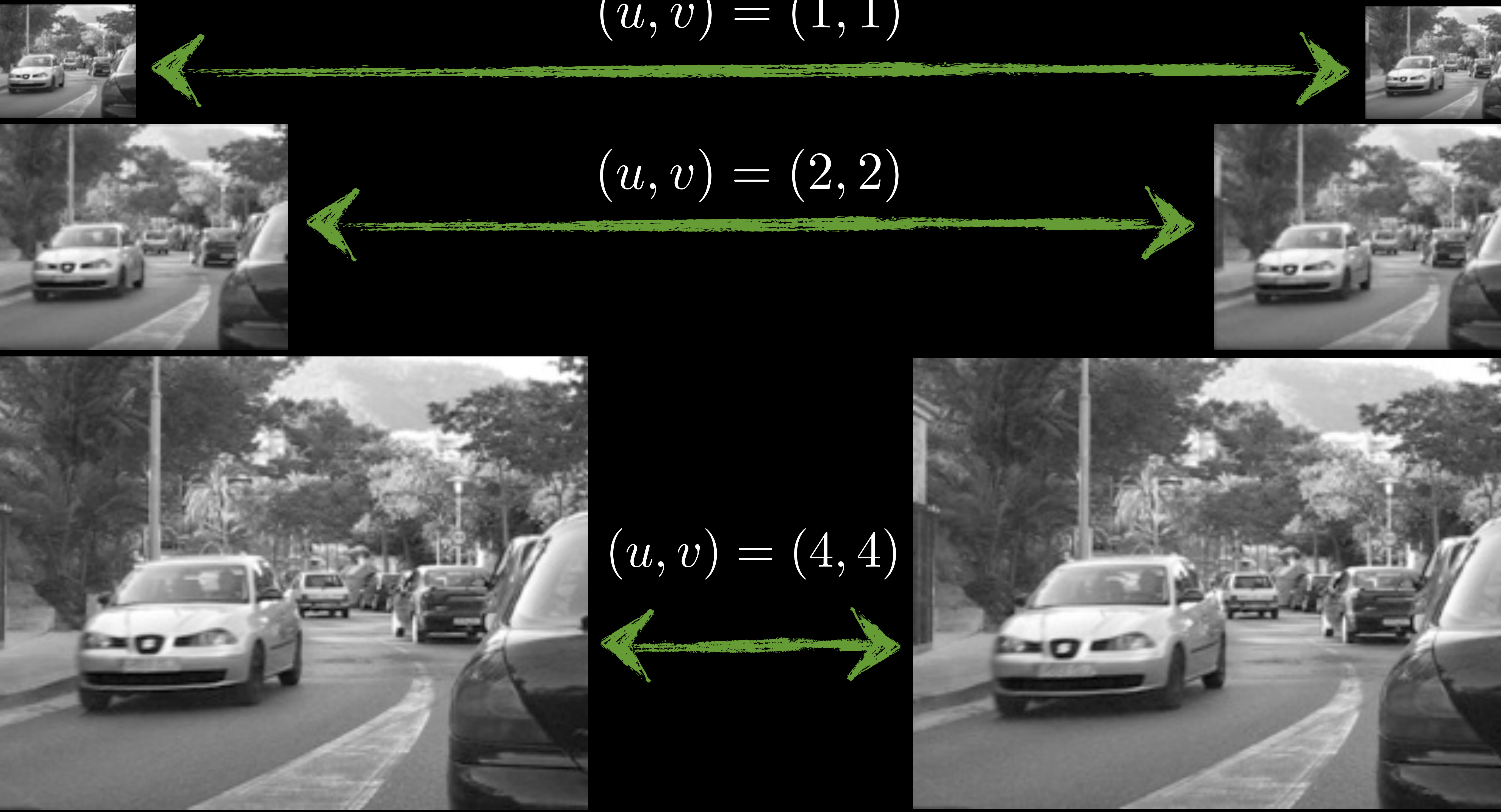
- Integrates information globally
- Dense flow

Cons

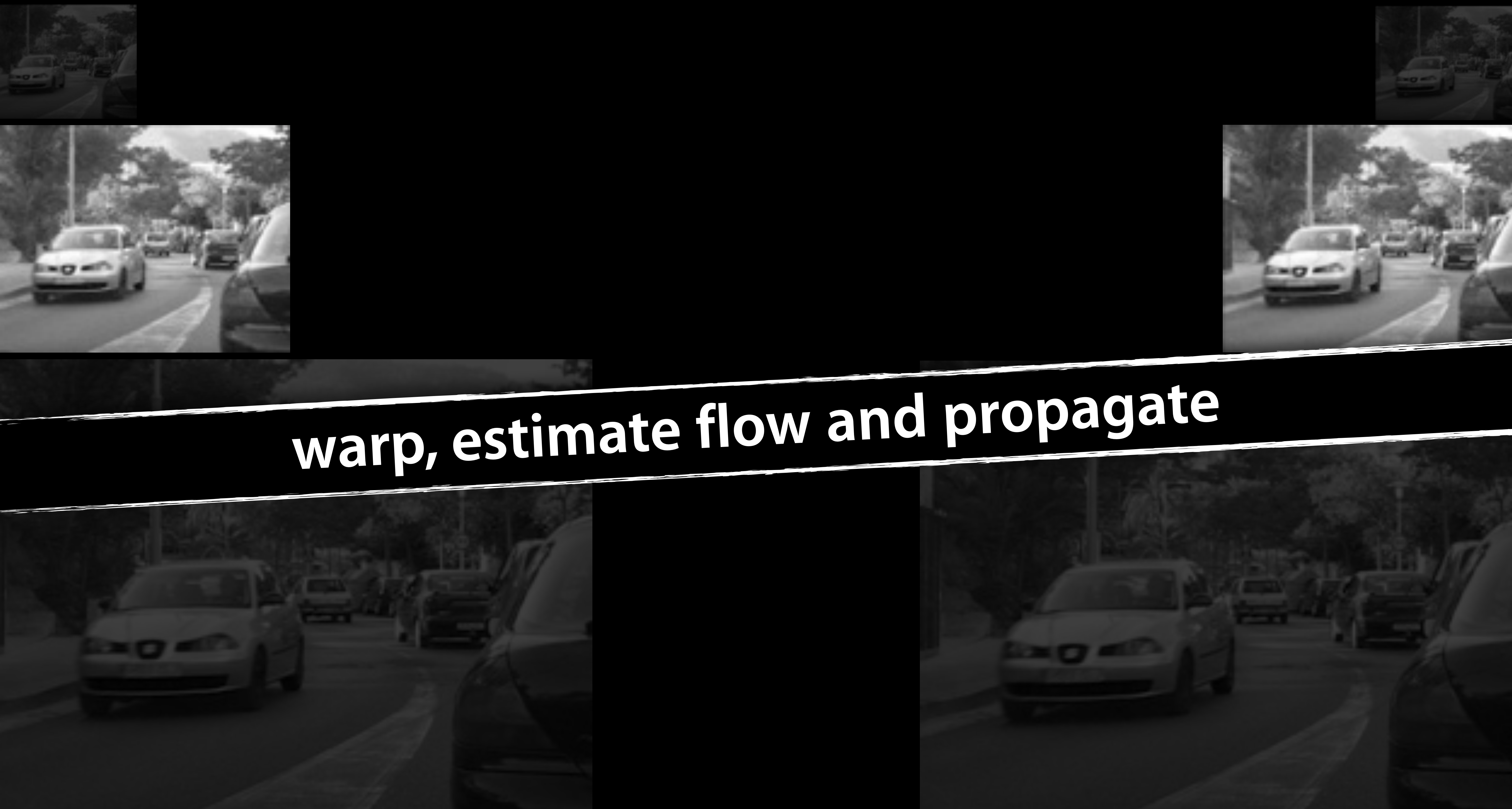
- Fails when brightness is not constant
- Slow iterative estimation
- Smooths over boundaries

When the velocity is large the input is **temporally aliased**

temporal derivatives not reliable



$I(x, y, t)$ $\quad$ $I(x, y, t + 1)$



$I(x, y, t)$

$I(x, y, t+1)$

Large displacements are difficult to estimate