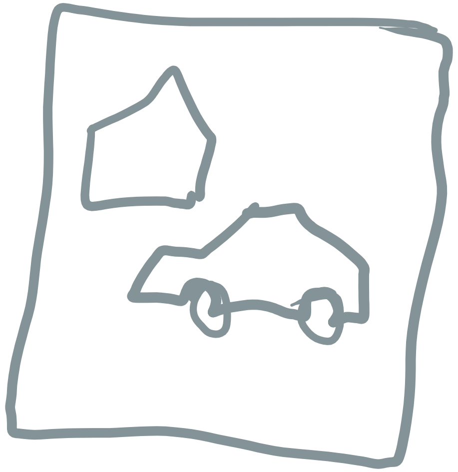
Motion estimation
Optical flow

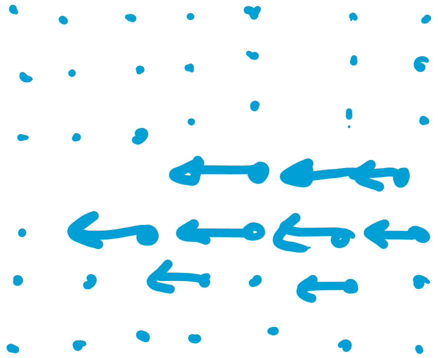
Computer Vision (SJK02)

Universitat Jaume I

Optic flow

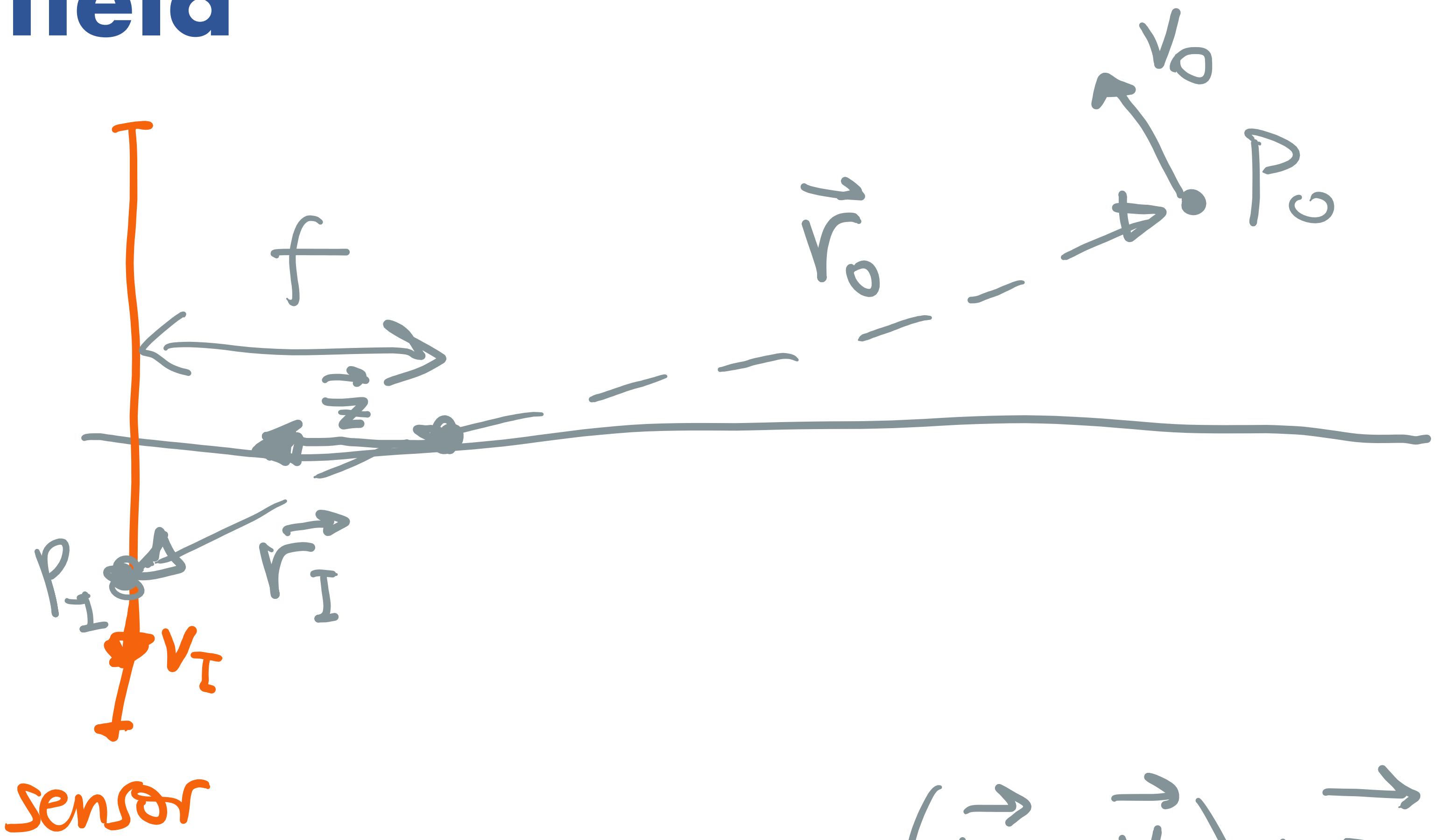
Goal: estimating *apparent* motion of objects in a scene from a sequence of images





[Optic flow. Overview (YouTube)]

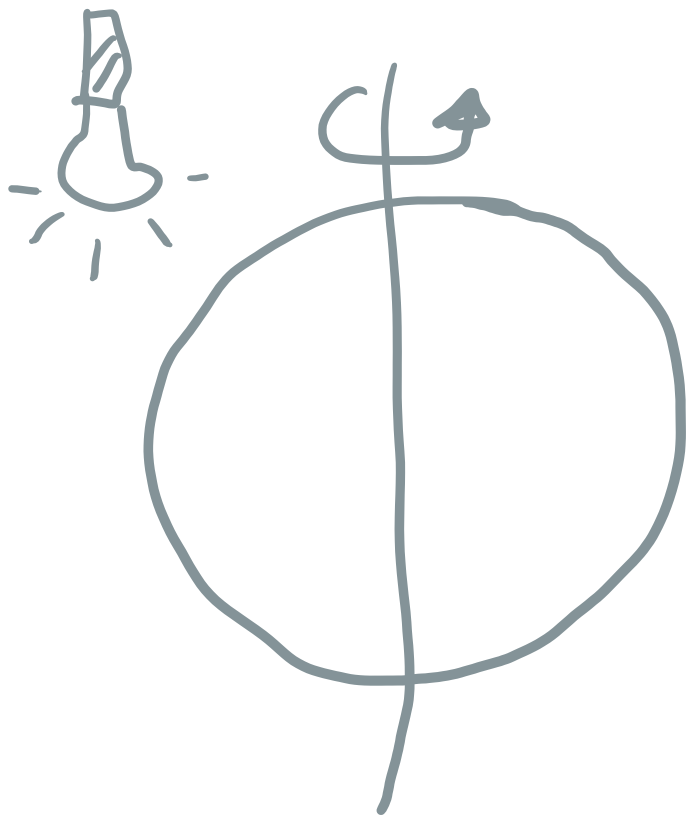
Motion field

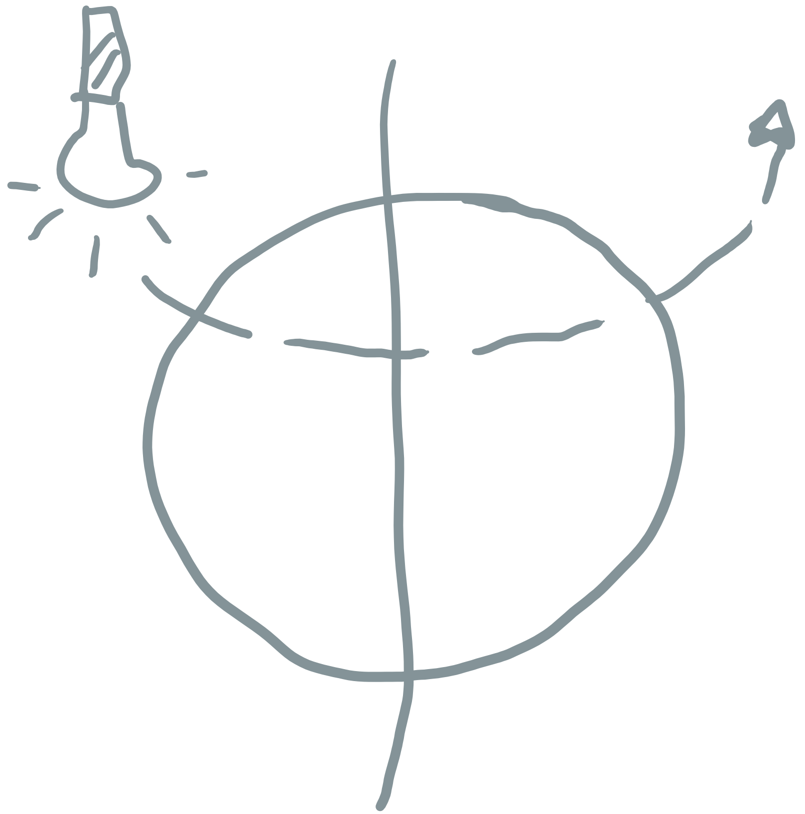


$$v_I = \frac{(\vec{r}_o \times \vec{v}_o) \times \vec{z}}{(\vec{r}_o \cdot \vec{z})^2}$$

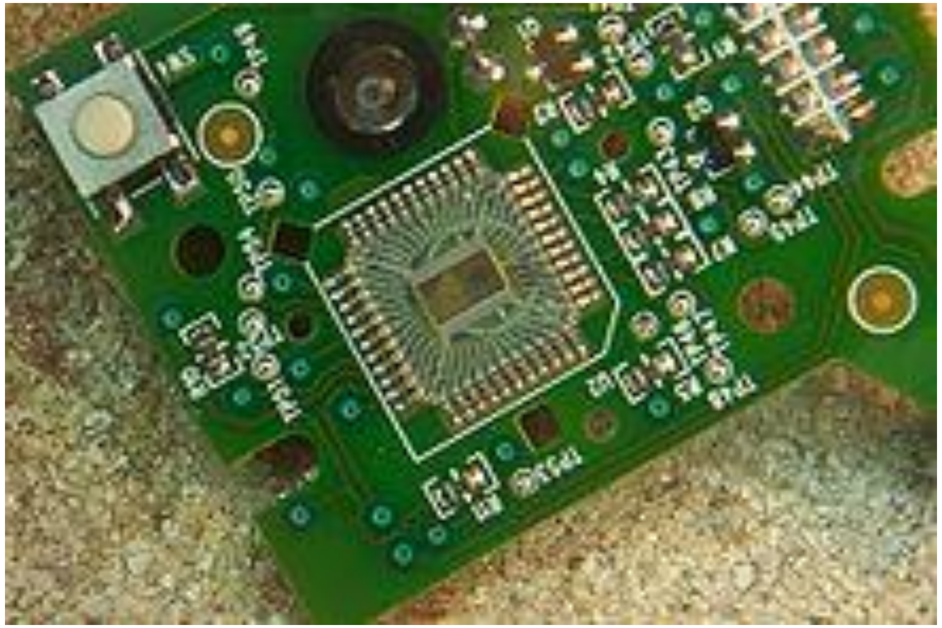
[Motion field vs optic flow (YouTube)](#)

Motion field vs Optic flow





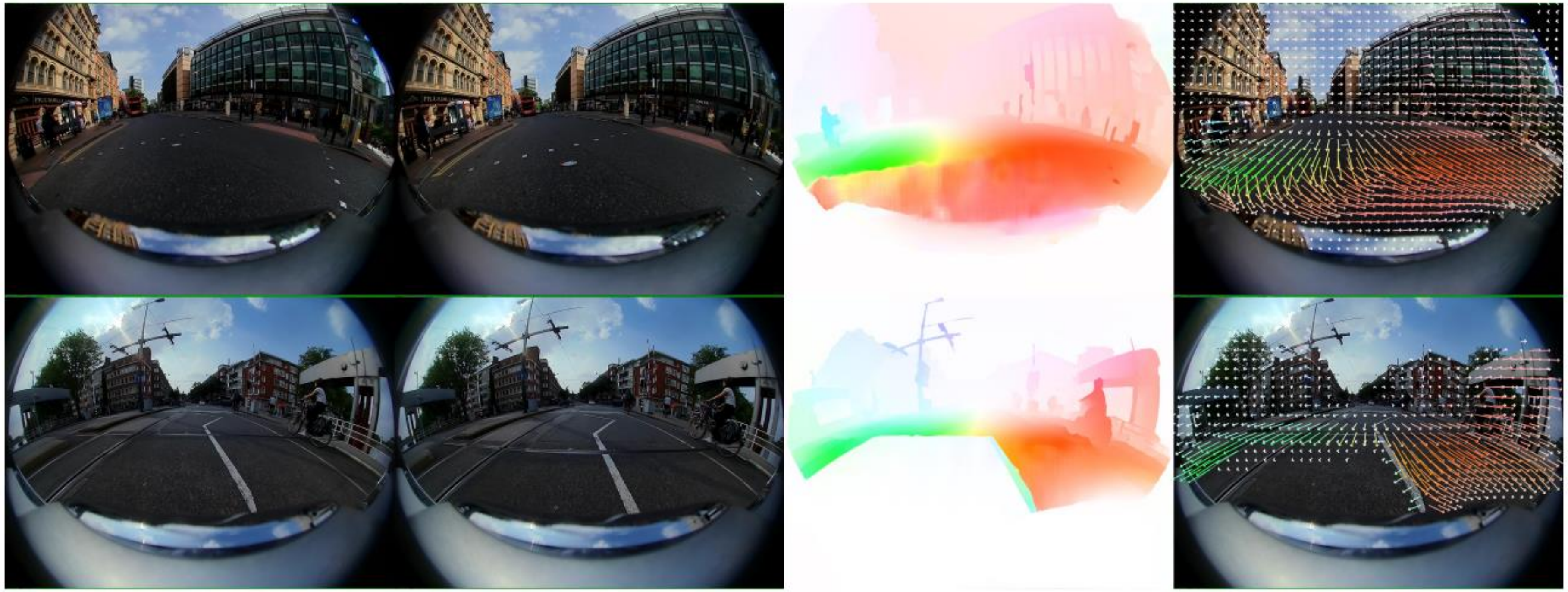
Application: Optical mouse



18 x 18
~ 1k frames/s

https://en.wikipedia.org/wiki/Optical_mouse

Application: Autonomous driving



object tracking
visual odometry
semantic segmentation
motion segmentation
SLAM

Optical Flow for Autonomous Driving: Applications, Challenges and Improvements (arXiv 2023)

Application: Action recognition

On the Integration of Optical Flow and Action Recognition (GCPR 2018)

- Optical flow is useful because it is **invariant to appearance**, **even** when the flow vectors are **inaccurate**
- EPE of current methods is **not well correlated** with action **recognition performance**
- Training optical flow to **minimize classification error** instead of minimizing EPE improves recognition performance

Application: video stabilisation

input
output

More videos at project page

SteadyFlow: Spatially Smooth Optical Flow for Video Stabilization (CVPR 2014)

Application: *Your turn*

Think of an (interesting) application of optical flow estimation

www.socrative.com
Room 219986

Theoretical assumptions

$$u = \frac{\delta x}{\delta t}, \quad v = \frac{\delta y}{\delta t}$$

(x, y) at t

(u, v)

$(x+\delta x, y+\delta y)$ at $t+\delta t$

Assumption 1: Brightness remains constant over time

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

Assumption 2: Displacement and time step are small

$I(x+\delta x, y+\delta y, t+\delta t)$ can be linearly approximated

[Optical flow constrained equation (YouTube)]

Taylor series expansion

$$f(x+\delta x) = f(x) + \frac{\partial f(x)}{\partial x} \cdot \frac{\delta x}{1!} + \cdots + \frac{\partial^{n} f(x)}{\partial x^{n}} \frac{(\delta x)^n}{n!}$$

If δx is small:

$$f(x+\delta x) = f(x) + \frac{\partial f(x)}{\partial x} \cdot \delta x + \text{HOT}$$

Linear approximation by ignoring H.O.T.

In our case, we have three variables (x, y, t)

$$I(x+\delta x,\ y+\delta y,\ t+\delta t) \cong I(x,y,t) + \frac{\partial I}{\partial x}\delta x + \frac{\partial I}{\partial y}\cdot\delta y + \frac{\partial I}{\partial t}\cdot\delta t$$

$$I(x+\delta x, y+\delta y, t+\delta t) = I(x,y,t) + I_x\delta x + I_y\delta y + I_t\delta t$$

Optical flow constraint equation

1 $I(x+\delta x, y+\delta y, t+\delta t) = I(x,y,t)$

2 $I(x+\delta x, y+\delta y, t+\delta t) = I(x,y,t) + I_x \delta x + I_y \delta y + I_t \delta t$

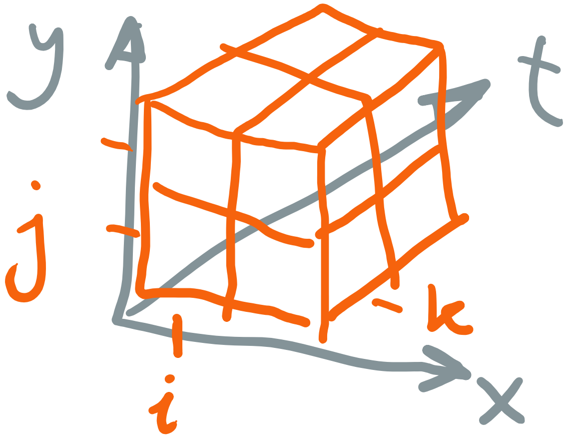
2 − 1

$$0 = I_x \delta x + I_y \delta y + I_t \delta_t$$

$$\lim_{\delta t \to 0} \quad I_x \frac{\delta x}{\delta t} + I_y \frac{\delta y}{\delta t} + I_t \frac{\delta t}{\delta t} = 0$$

$$I_x \frac{\partial x}{\partial t} + I_y \frac{\partial y}{\partial t} + I_t = 0 \qquad I_x u + I_y v + I_t = 0$$

Computing partial derivatives



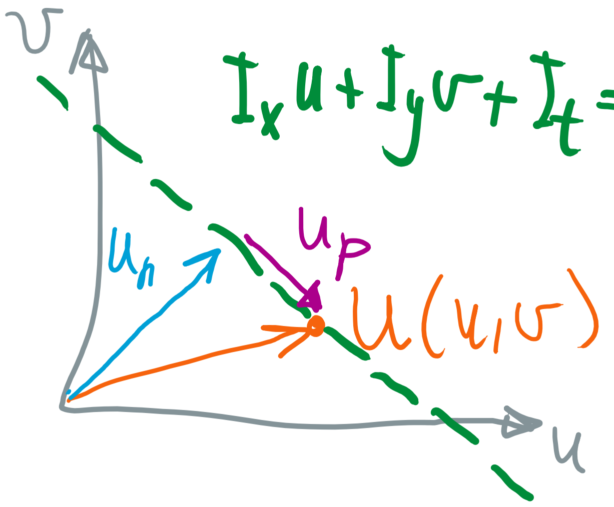
$I_x(i,j,k) =$

$I_y(i,j,k) =$

$I_t(i,j,k) =$

Geometric interpretation of OF equation

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



underconstrained

$$u = u_n + u_p$$

$$r \equiv ax + by + c = 0$$

$$\vec{n} = (a, b)$$

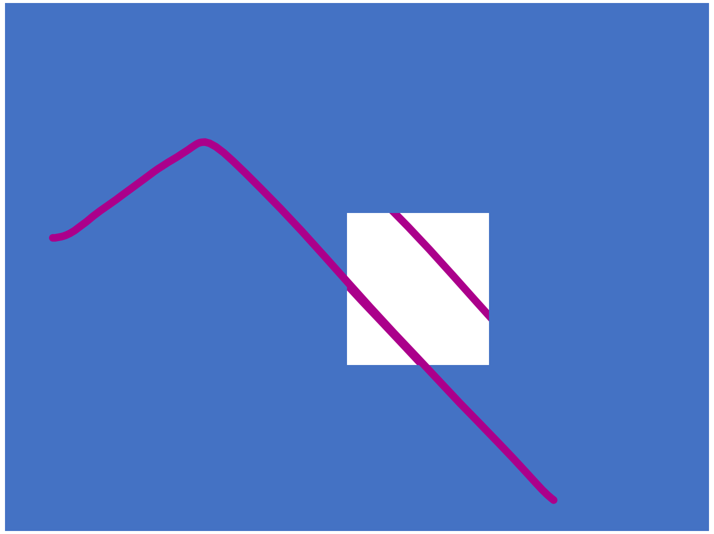
$$d(P(x_0, y_0), r) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$

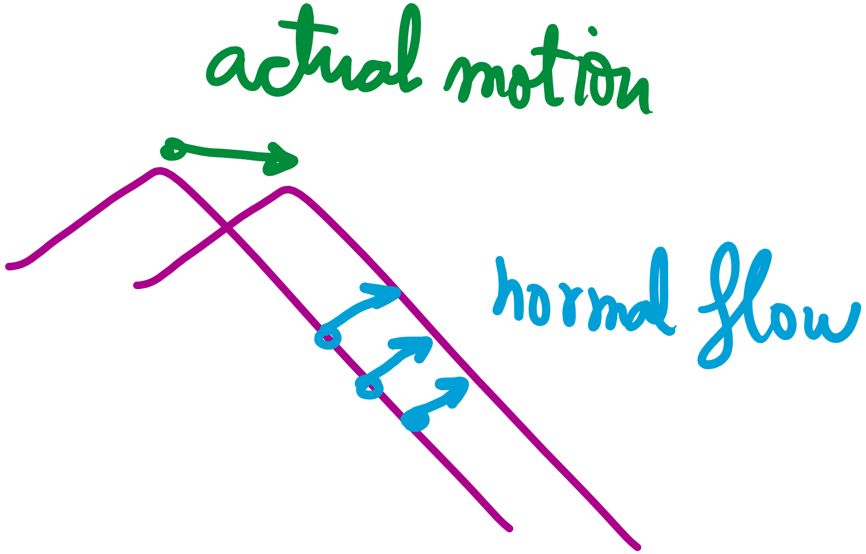
$$\hat{u}_n = \frac{(I_x, I_y)}{\sqrt{I_x^2 + I_y^2}} \qquad |u_n| = \frac{|I_t|}{\sqrt{I_x^2 + I_y^2}}$$

$$u_n = |u_n| \cdot \hat{u}_n = |I_t| \frac{(I_x, I_y)}{I_x^2 + I_y^2}$$

u_p?

Aperture problem





Locally, only the **normal** flow can be determined

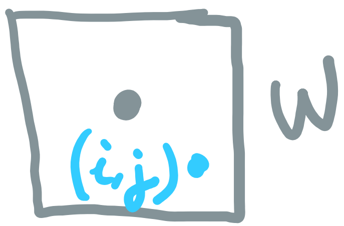
To fully solve the OF equation, we need additional constraints

True or false?
The aperture problem does not depend on the image contents; it is an intrinsic limitation of the optical flow constraint equation

www.socrative.com
Room 219986

Lucas-Kanade method

Assumption: flow is the same in a small neighbourhood *W*



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

$$I_x(i,j)\,u + I_y(i,j)\,v + I_t(i,j) = 0 \qquad \forall (i,j) \in W$$

$$\begin{bmatrix} I_x(1,1) & I_y(1,1) \\ I_x(1,2) & I_y(1,2) \\ \vdots & \vdots \\ I_x(n,n) & I_y(n,n) \end{bmatrix} \cdot \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(1,1) \\ I_t(1,2) \\ \vdots \\ I_t(n,n) \end{bmatrix}$$

$$Au = b$$

n equations,
2 unknowns
Linearly independent?

[Lucas-Kanade (YouTube)](#)

Least Squares

$$Au = b$$

$$A^T A u = A^T b$$

$$u = (A^T A)^{-1} A^T b$$

$$A^T A =$$

$$A^T b =$$

$$\underbrace{\begin{bmatrix} I_x(1,1) & I_y(1,1) \\ I_x(1,2) & I_y(1,2) \\ \vdots & \vdots \\ I_x(n,n) & I_y(n,n) \end{bmatrix}}_{A} \qquad \underbrace{-\begin{bmatrix} I_t(1,1) \\ I_t(1,2) \\ \vdots \\ I_t(n,n) \end{bmatrix}}_{b}$$

$$u = (A^T A)^{-1} A^T b$$

$A^T A$ Must be invertible

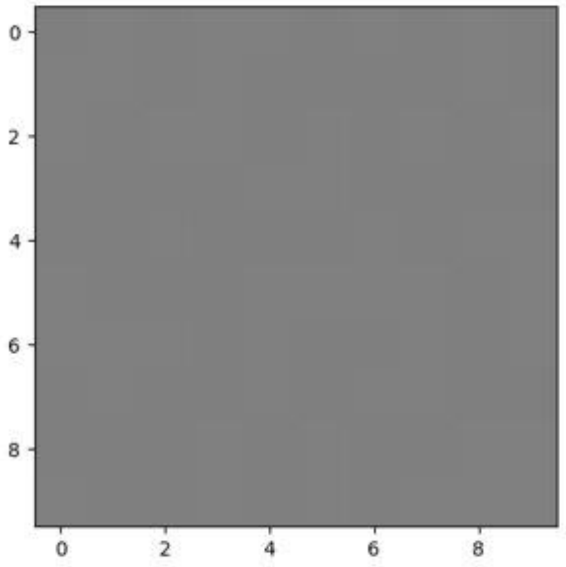
$A^T A$ Must be well-conditioned

$\lambda_1, \lambda_2 \equiv$ eigen values of $A^T A$

$\lambda_1 > \varepsilon, \ \lambda_2 > \varepsilon$

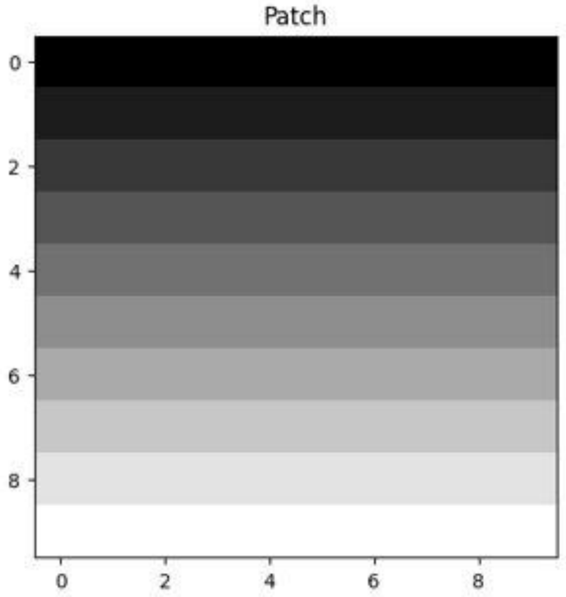
$\lambda_1 > \lambda_2$ (but not $\lambda_1 \gg \lambda_2$)

textureless



$\lambda_1, \lambda_2 \simeq 0$

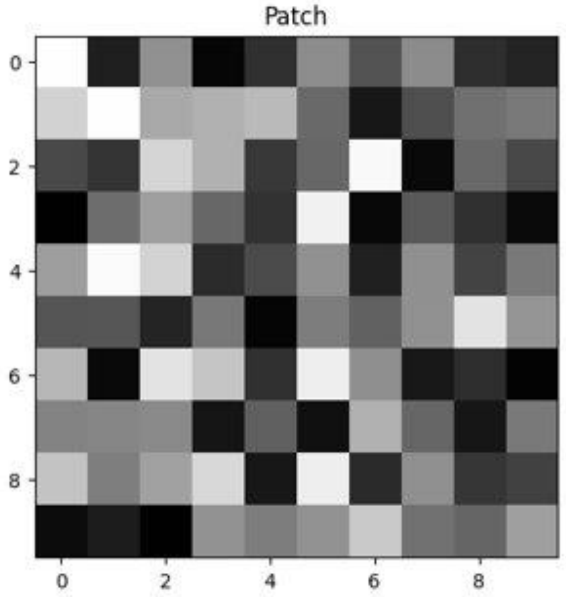
edge



$\lambda_1 \gg \lambda_2$

(aperture!)

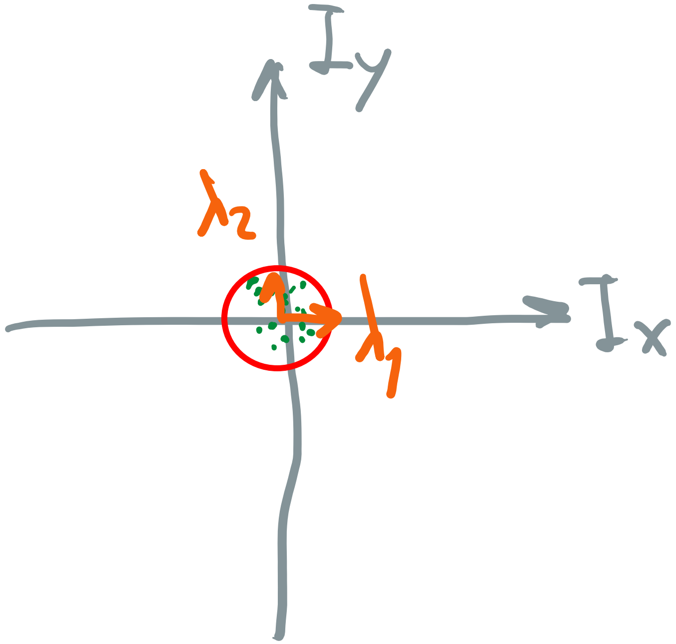
textured



$\lambda_1 \approx \lambda_2$

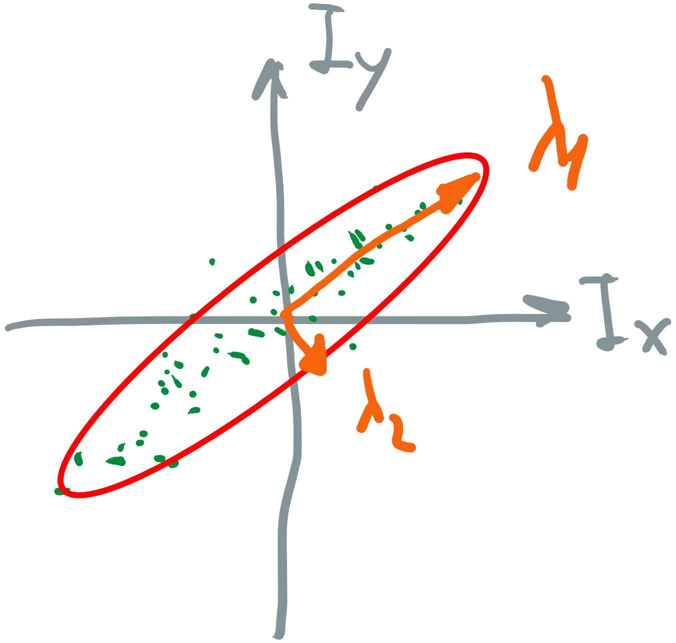
$\lambda_1, \lambda_2 > \varepsilon$

textureless



$\lambda_1, \lambda_2 \simeq 0$

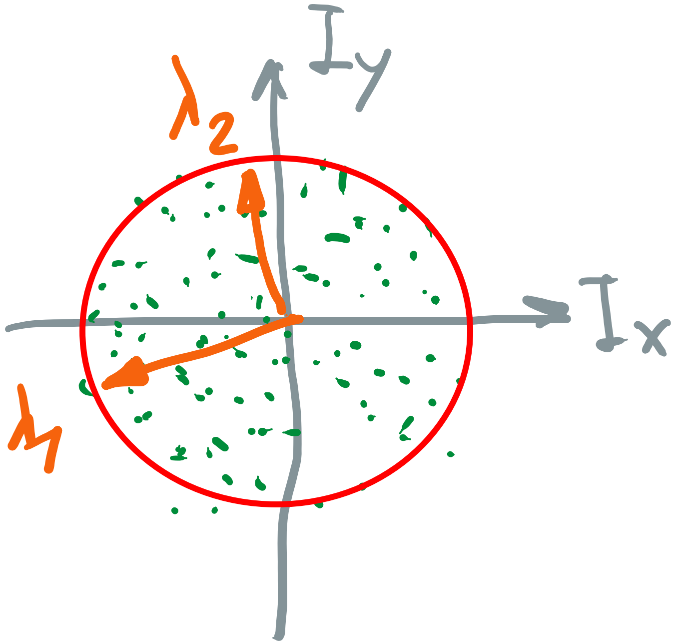
edge



$\lambda_1 \gg \lambda_2$

(aperture!)

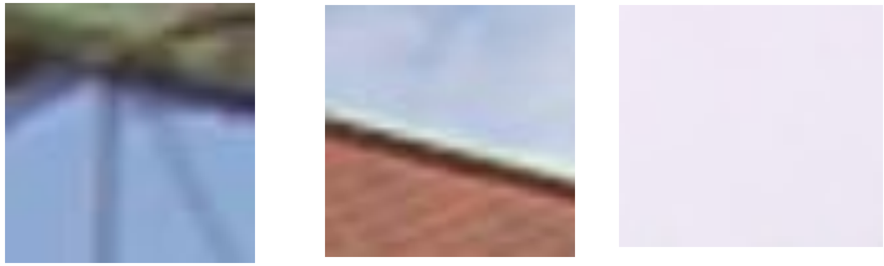
textured



$\lambda_1 \approx \lambda_2$

$\lambda_1, \lambda_2 > \varepsilon$

A B C



Where can OF estimates be more reliable?

www.socrative.com

Room 219986

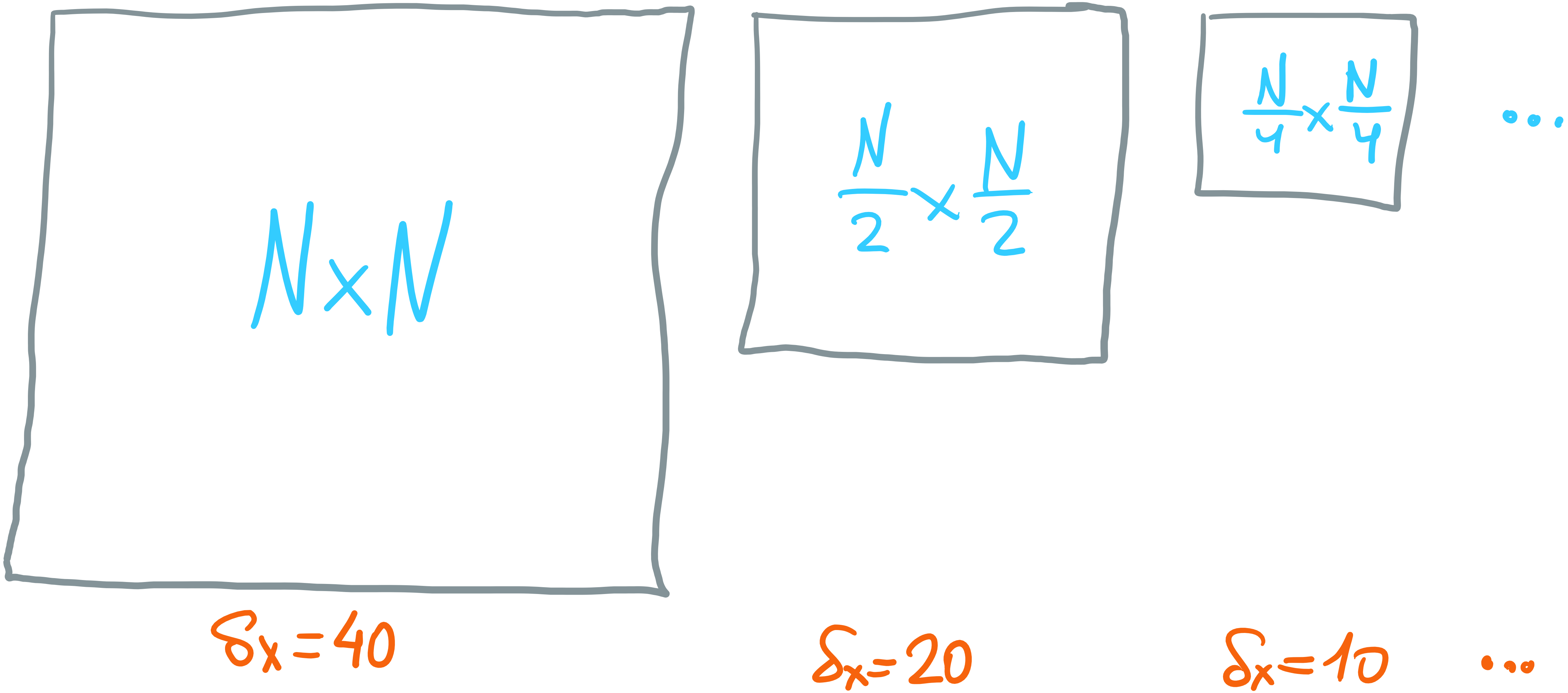
What about large motions?

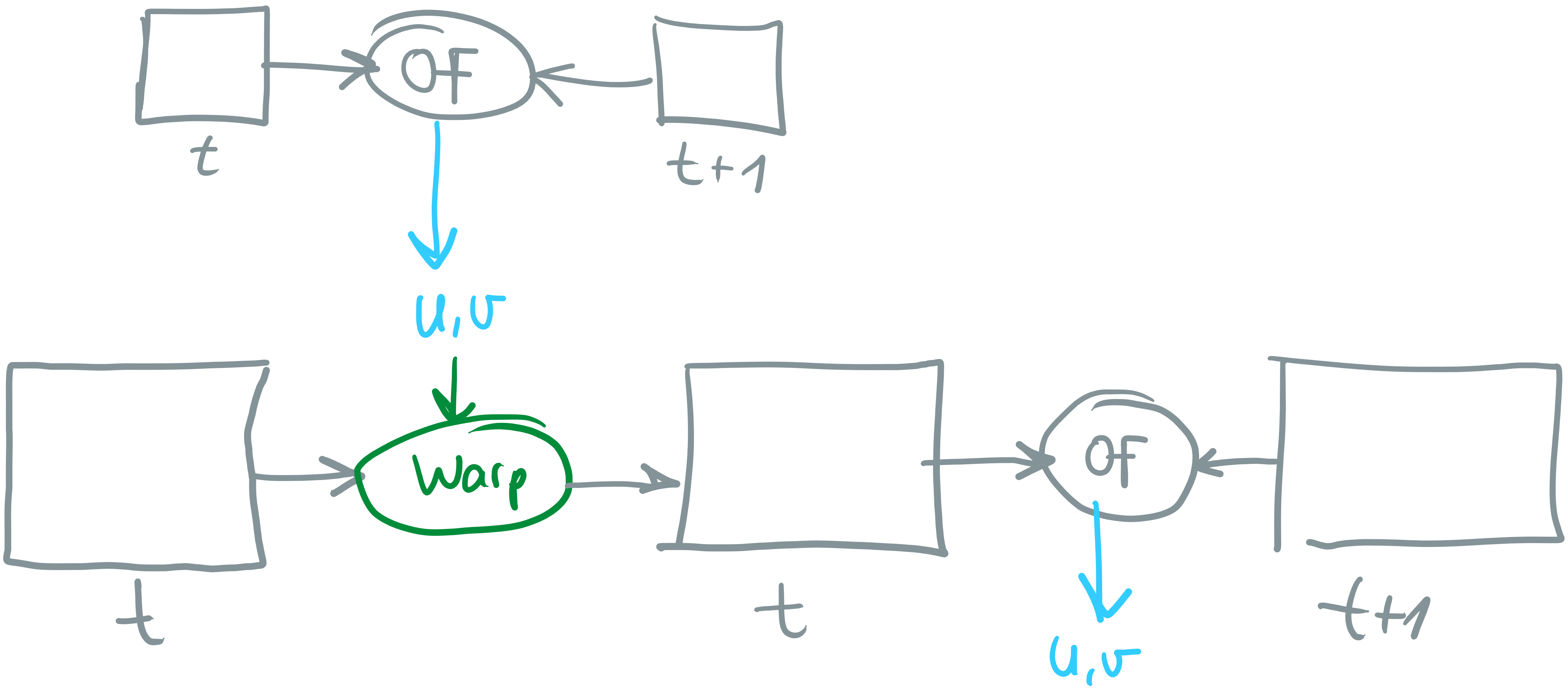
Linear approximation no longer valid

Optic flow equation not valid

[Coarse-to-fine approach (YouTube)]

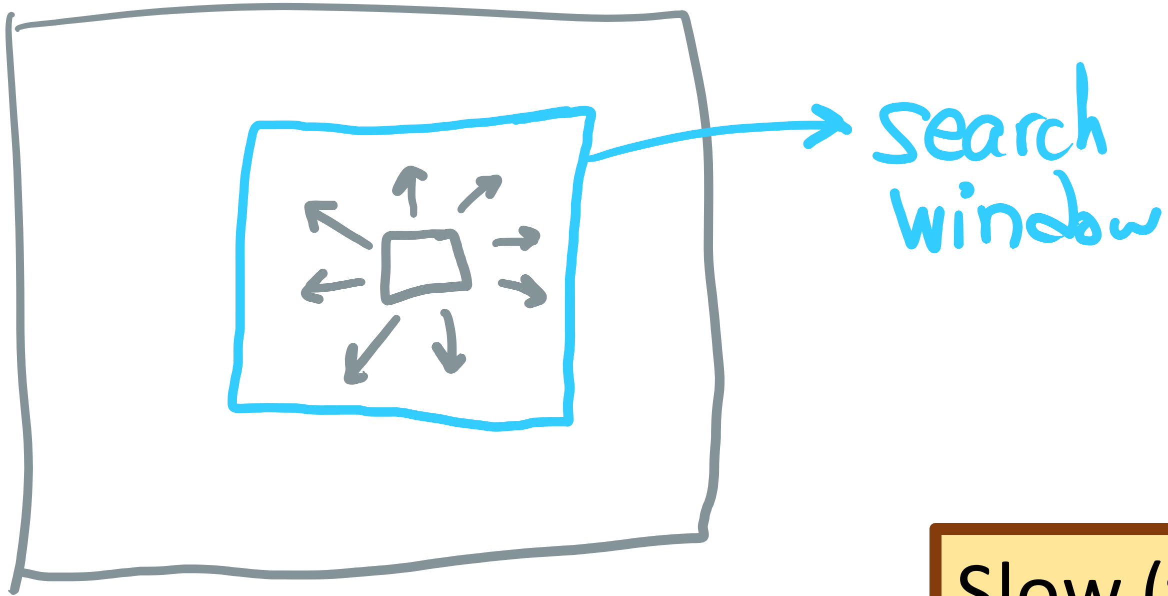
Coarse-to-fine strategy





Repeat OF and warp (accumulating flows!)

Alternative approach: template matching



Slow (for large search windows)

Mismatches are possible

Horn-Schunk method

$$E = \iint \left[(I_x u + I_y v + I_t)^2 \qquad\qquad\qquad\qquad \right] \mathrm{d}x\mathrm{d}y$$

Smoothness term

Would we minimise or maximise *E*?

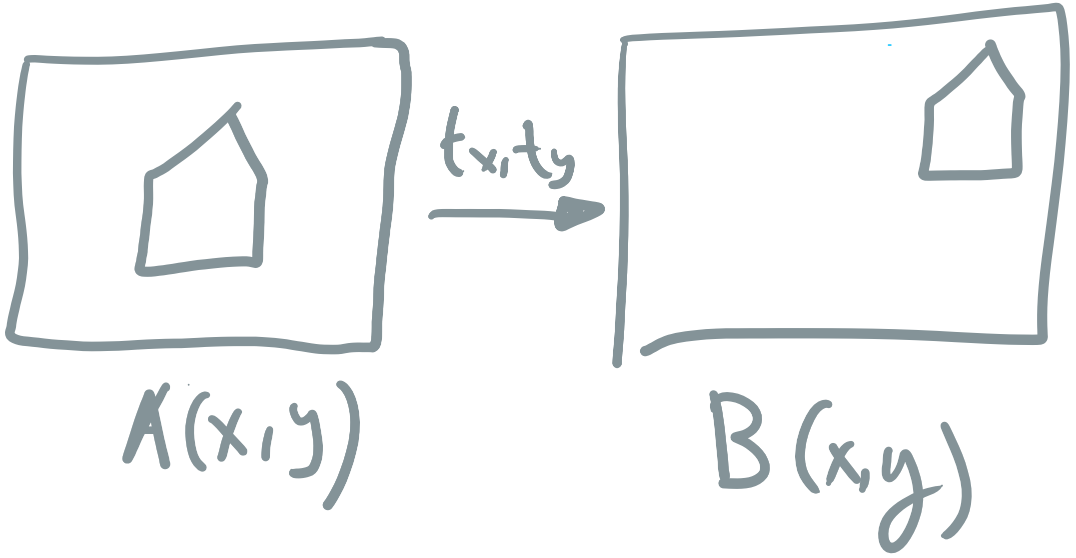
Dense or sparse method?

[Horn-Schunck Optical Flow with a Multi-Scale Strategy (demo)](https://en.wikipedia.org/wiki/Horn-Schunck_method)

https://en.wikipedia.org/wiki/Horn-Schunck_method

Global motion
Image registration

Transformation model



$$T(x, y; t_x, t_y) = (x + t_x, y + t_y)$$

$$B(x, y) = A\big(T(x, y; t_x, t_y)\big)$$

θ = set of motion parameters

$$T(x, y; \theta)$$

translation: $\theta = (t_x, t_y)$

rotation: $\theta = (\beta)$

scale: $\theta = (\alpha_x, \alpha_y)$

affine: $\theta = (a_x, b_x, t_x, a_y, b_y, t_y)$

Projective $\theta = (\ldots$ 8 params $\ldots)$

...

Registration problem

$$\theta^* = \arg\underset{\theta}{\mathrm{opt}}\ f(A, B, T_\theta)$$

SSD (Sum of Squared Differences)

$$\theta^* = \arg\min_{\theta} \sum_{x,y} \left[A(x,y) - B\left(T_\theta(x,y)\right)\right]^2$$

Can we use SSD if A and B have different ranges of gray values?

Some demos

[Homography & RANSAC]

[Coarse-to-fine, least squares]