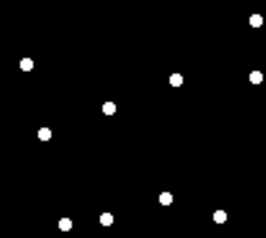
Introduction to motion correspondence



IPAM - UCLA

July 24, 2013

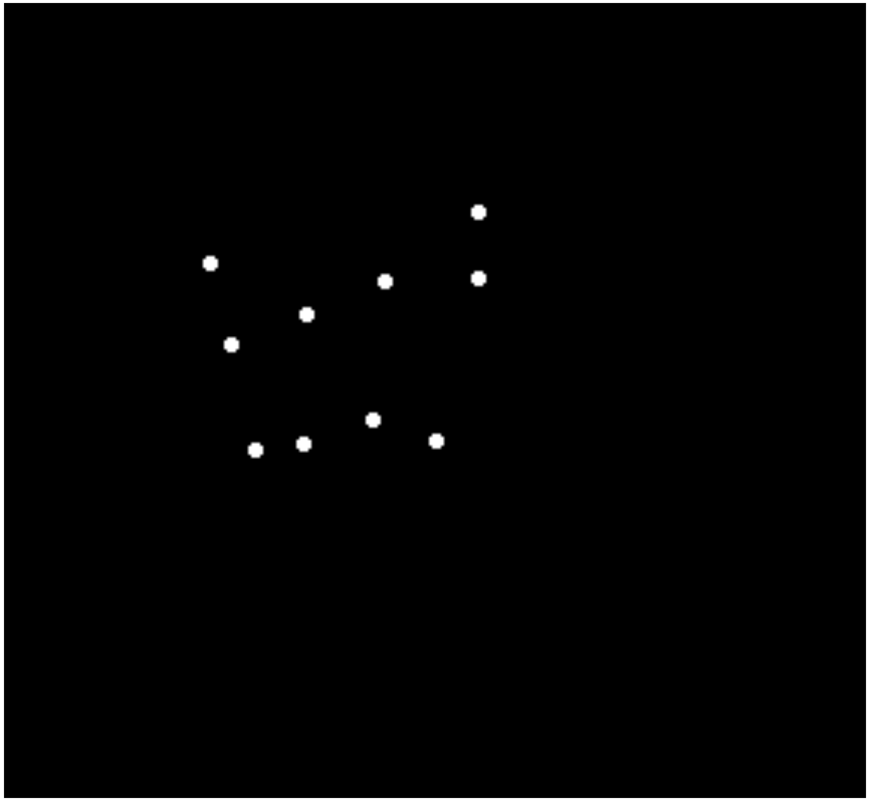
Iasonas Kokkinos
Center for Visual Computing
Ecole Centrale Paris / INRIA Saclay

Why estimate visual motion?

- Tracking
- Segmentation
- Structure from motion
- Action recognition

Action recognition and motion

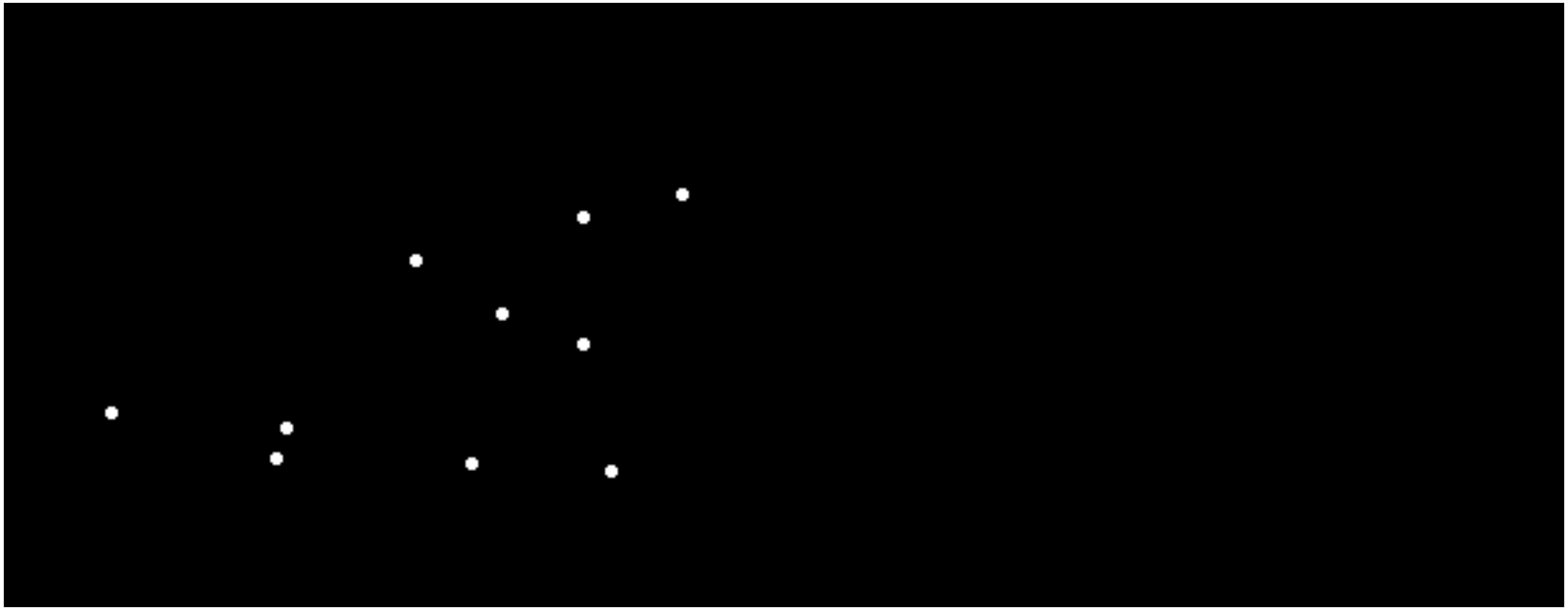
- http://astro.temple.edu/~tshipley/mocap/dotMovie.html



G. Johansson, "Visual Perception of Biological Motion and a Model For Its Analysis", *Perception and Psychophysics 14, 201-211, 1973.*

Action recognition and motion

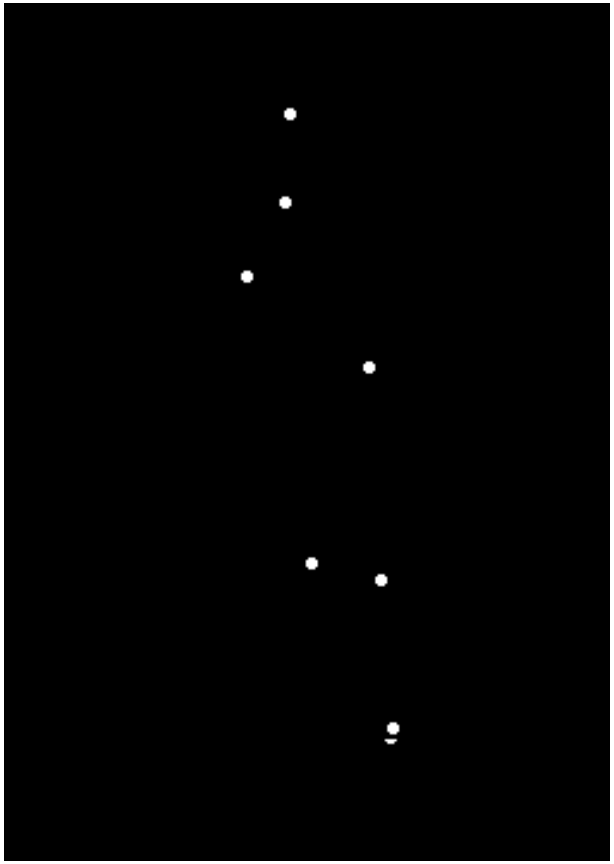
- http://astro.temple.edu/~tshipley/mocap/dotMovie.html



G. Johansson, “Visual Perception of Biological Motion and a Model For Its Analysis", *Perception and Psychophysics 14, 201-211, 1973.*

Action recognition and motion

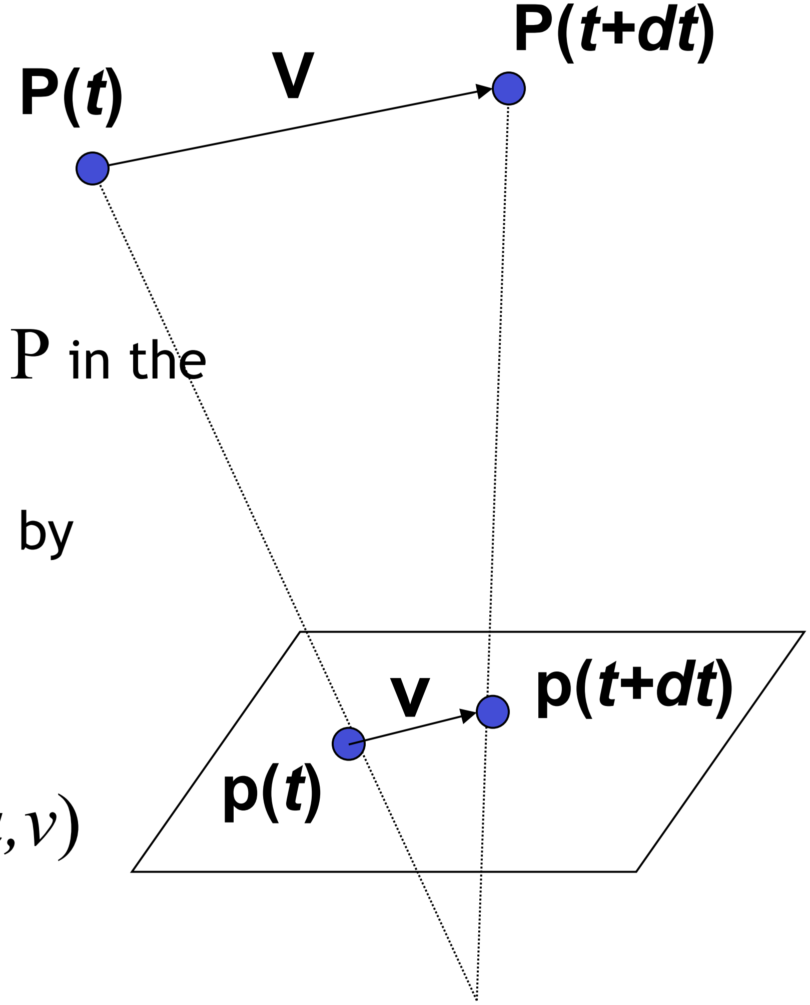
- http://astro.temple.edu/~tshipley/mocap/dotMovie.html



G. Johansson, “Visual Perception of Biological Motion and a Model For Its Analysis", *Perception and Psychophysics 14, 201-211, 1973.*

Motion Field

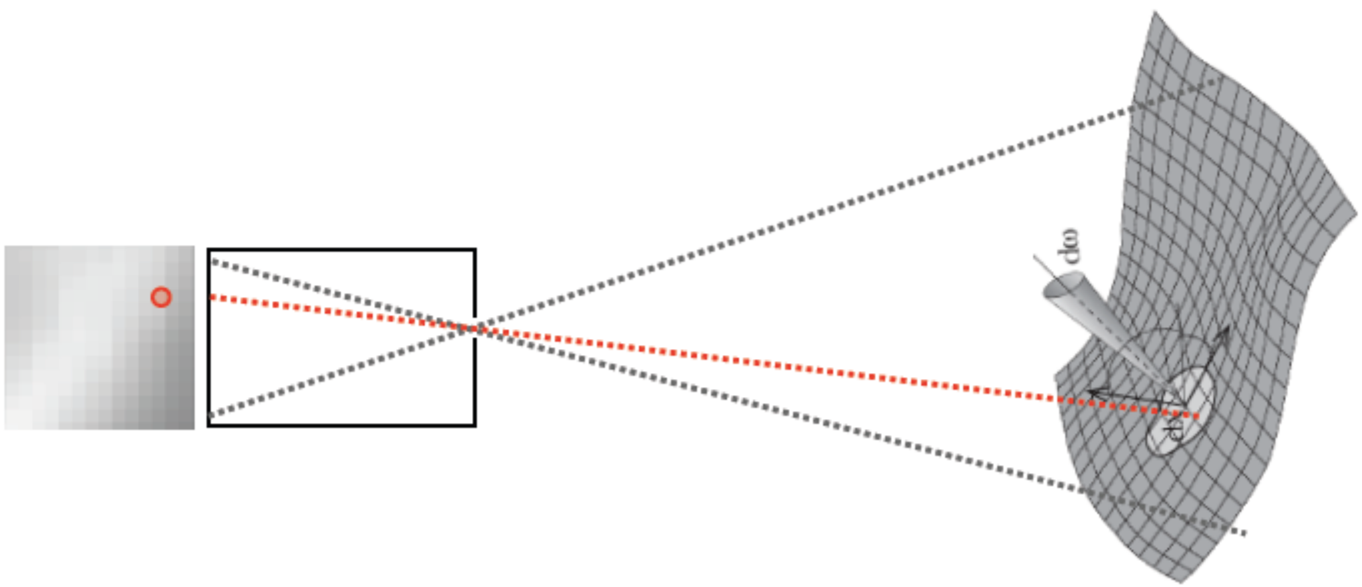
- $\mathrm{P}(t)$ is a moving 3D point
- Velocity of scene point:
  $\mathrm{V} = \mathrm{dP}/\mathrm{d}t$
- $\mathrm{p}(t) = (x(t),y(t))$ is the projection of P in the image.
- Apparent velocity v in the image: given by components
  $u = \mathrm{d}x/\mathrm{d}t$ and $v = \mathrm{d}y/\mathrm{d}t$
- Considering $\mathrm{d}t = 1$: $\mathrm{p}(t+1) = \mathrm{p}(t) + (u,v)$
- Motion estimation task:
  - Estimate (u,v)



Slide credit: S. Lazebnik

Brightness constancy constraint

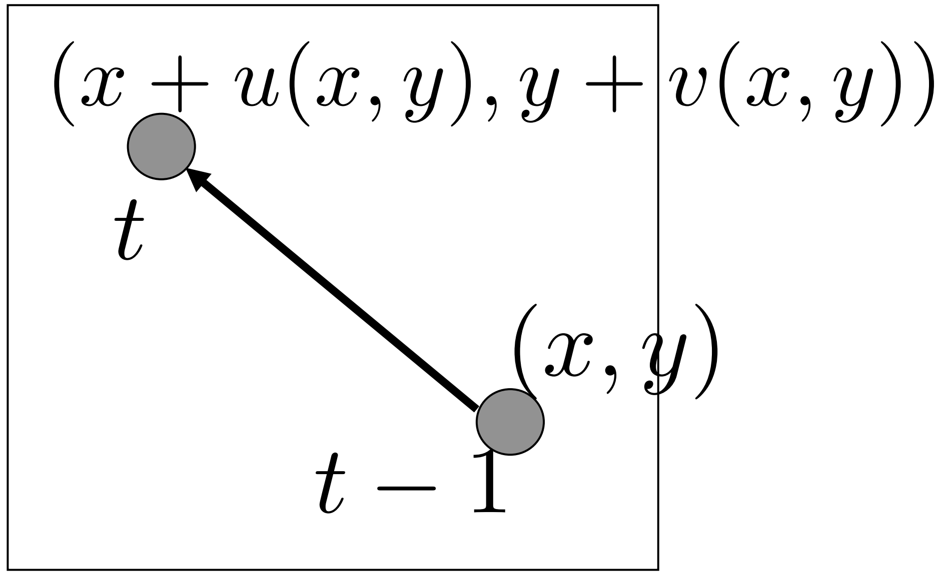
- Image is projection of 3D environment
- Each pixel: projection of a surface patch
- Pixel intensity influenced by 3D surface, incident light, camera …
- Assumption: intensity of surface patch remains constant



Slide credit: Georg Langs

Brightness constancy constraint

Optical flow estimation:



Constraint:

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

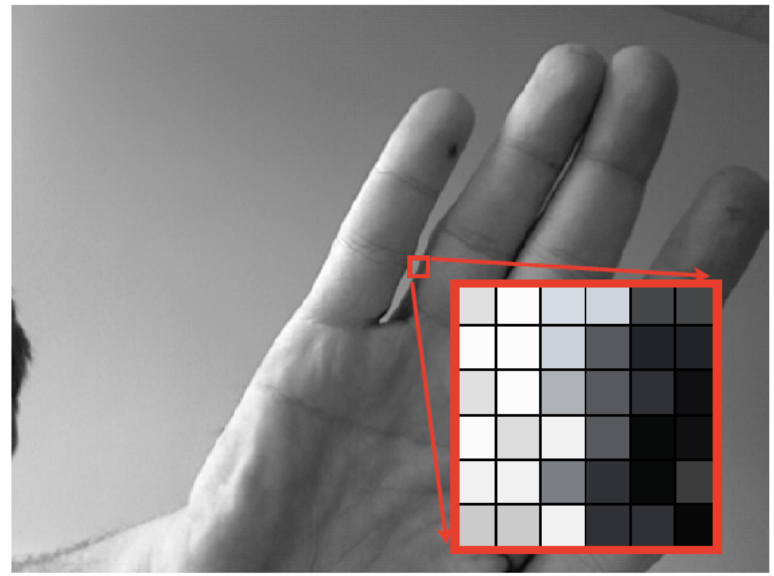
Taylor expansion of RHS:

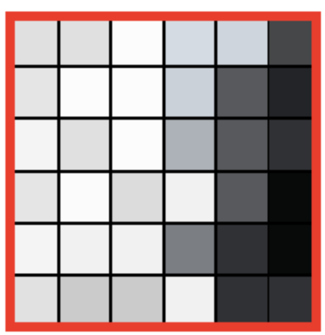
$$I(x,y,t-1) \simeq I(x,y,t) + \frac{\partial I}{\partial x} u(x,y) + \frac{\partial I}{\partial y} v(x,y)$$

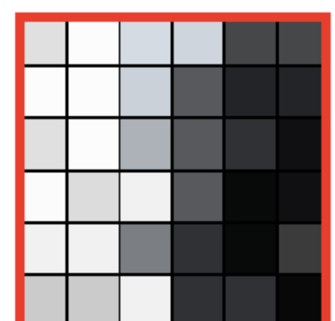
Brightness Constancy Constraint:

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

Optical flow estimation - 2D to 1D



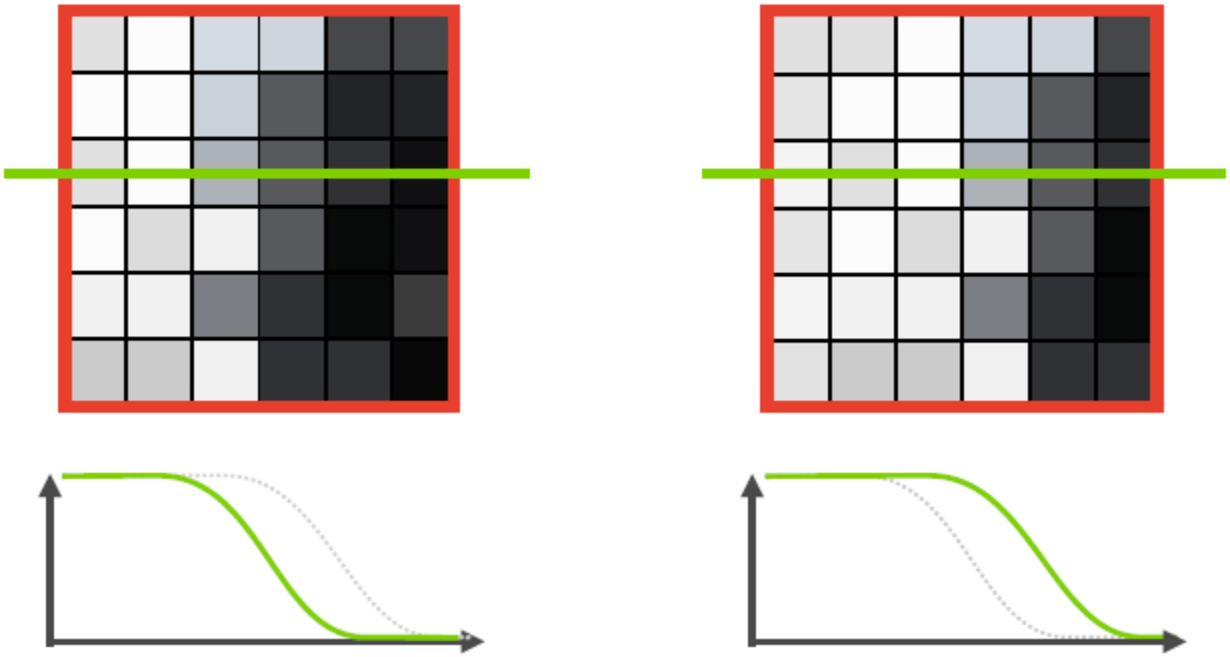




Slide credit: Georg Langs

Optical flow estimation - 1D case

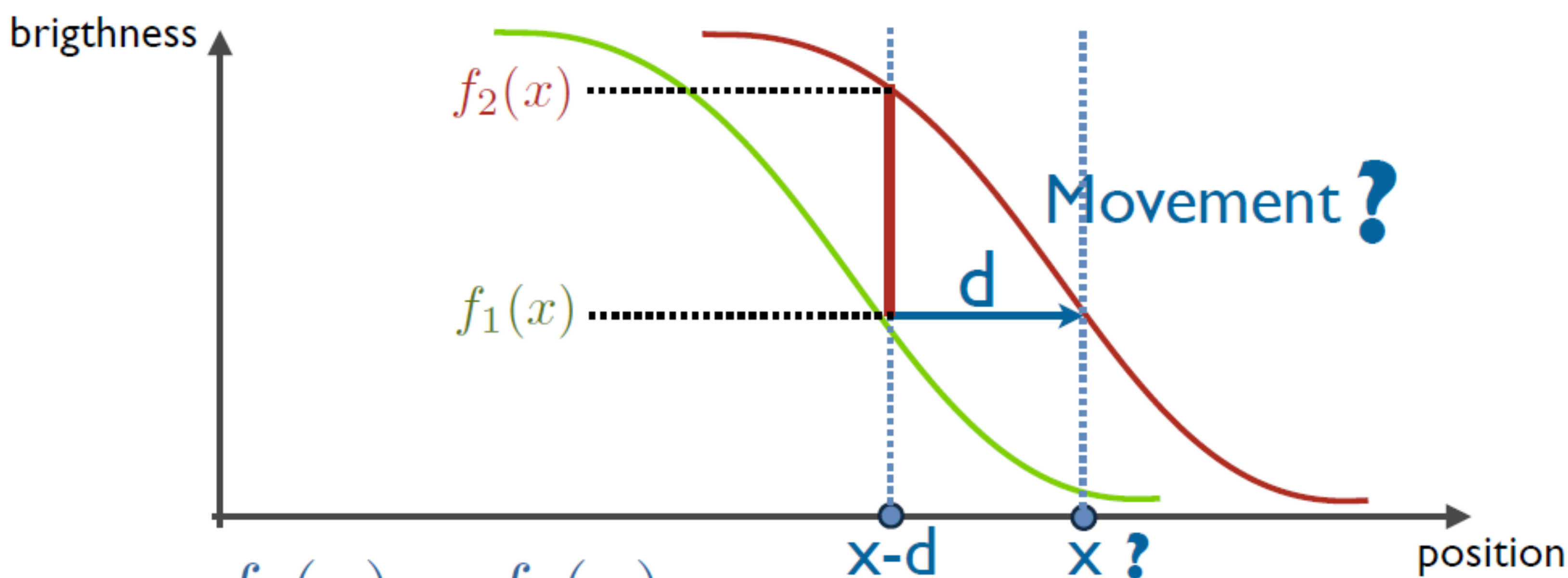
- How can we estimate the displacement?



Slide credit: Georg Langs

Optical Flow Estimation - 1D case

- Known: Gradient, Difference of intensities at x-d
- Unknown: d



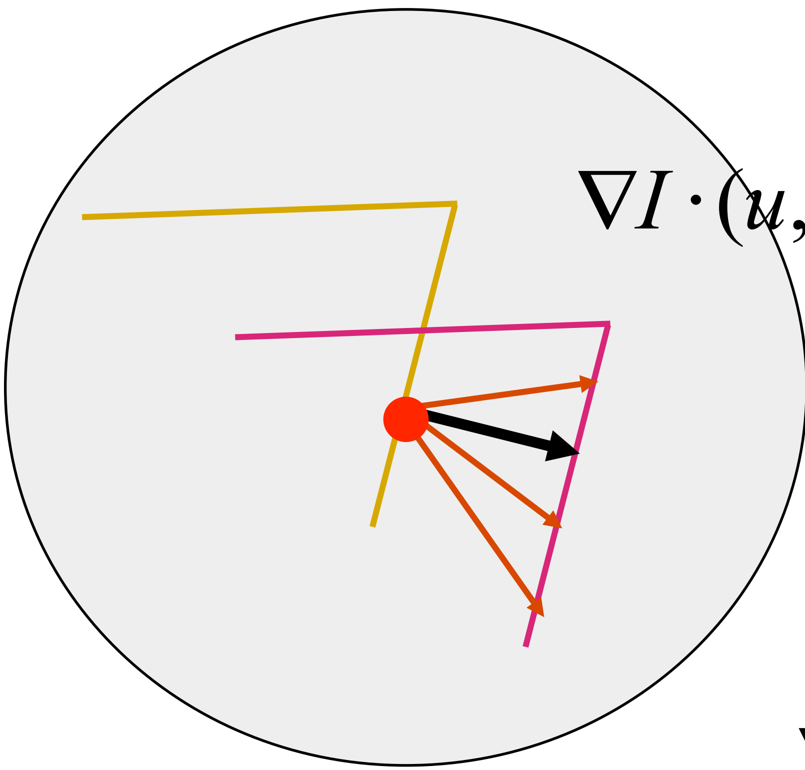
$$d = \frac{f_1(x) - f_2(x)}{f_1'(x)}$$

How about 2D?

Slide credit: Georg Langs

Brightness constraint: not enough for 2D!

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



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

$$(u', v') = (u, v) + c(-I_y, I_x)$$

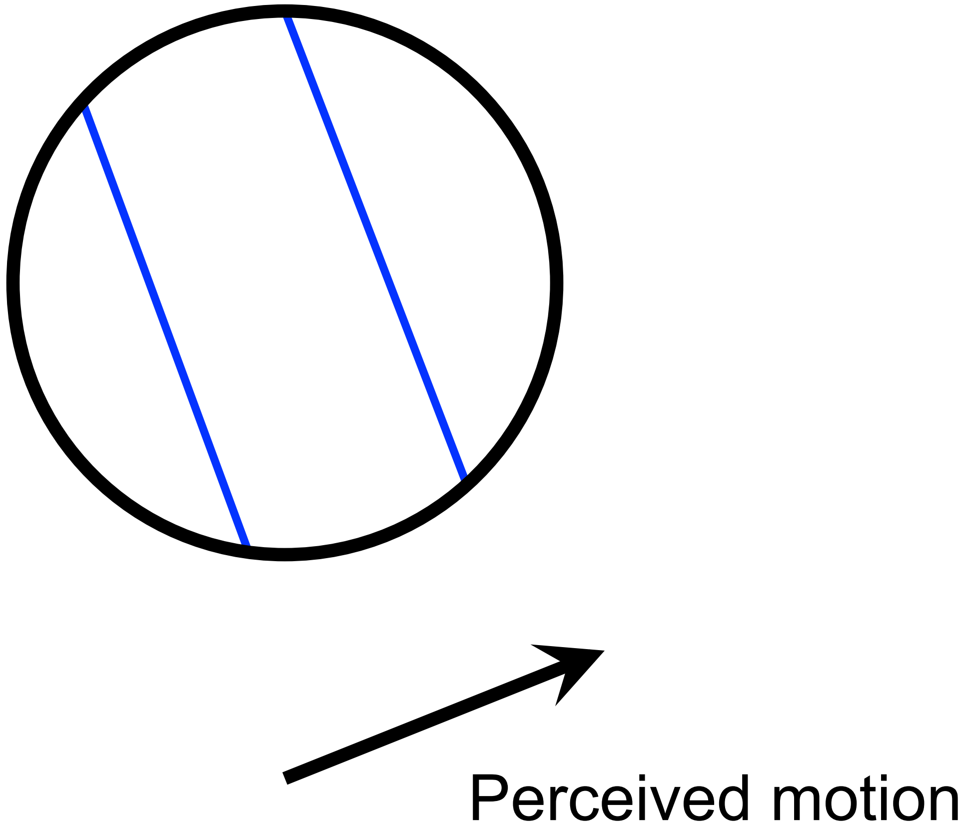
$$\nabla I \cdot (u, v) + I_t = 0$$

unknown flow parallel to the edge

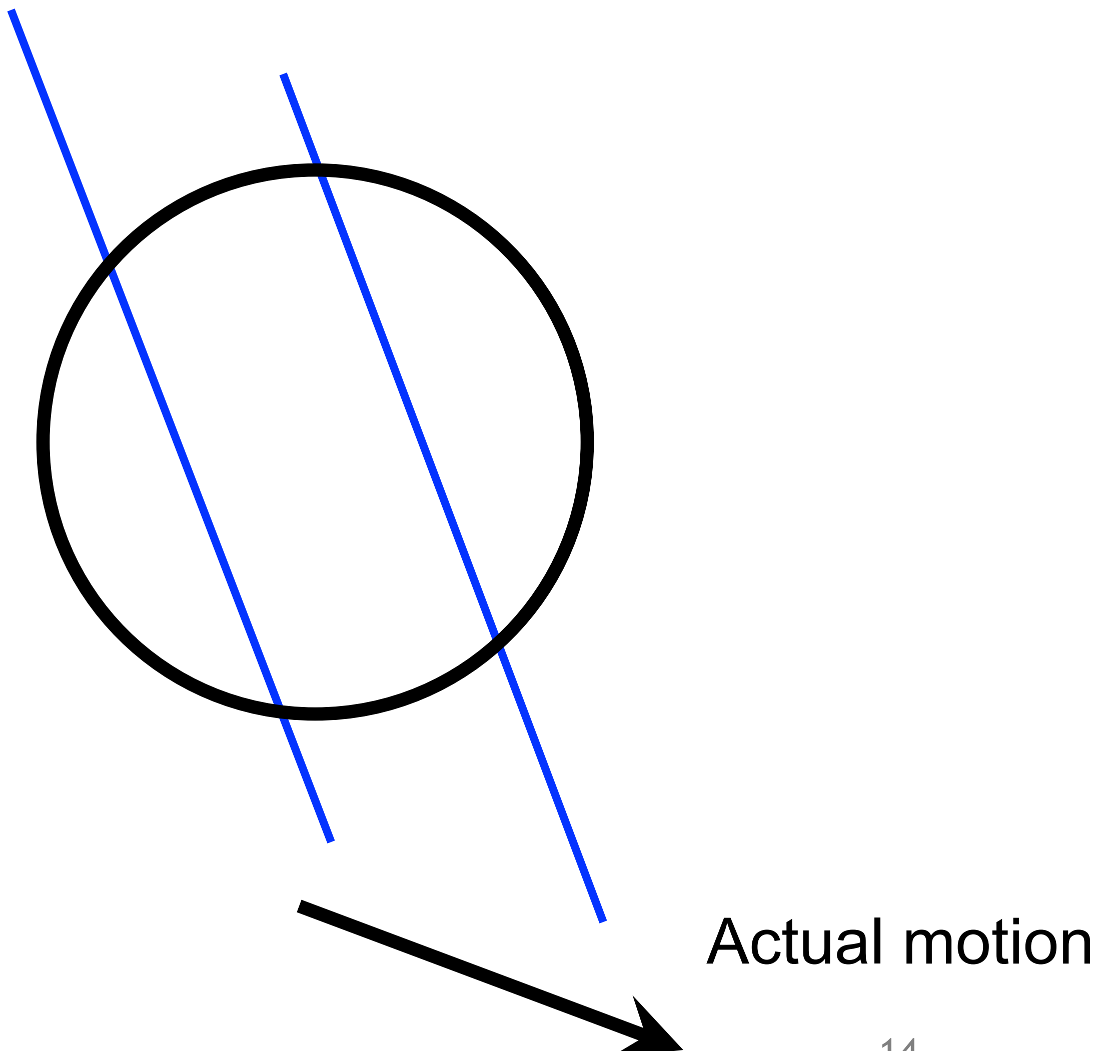
$$\nabla I \cdot (u, v) + I_t = 0$$

$$\nabla I \cdot (u', v') = 0$$

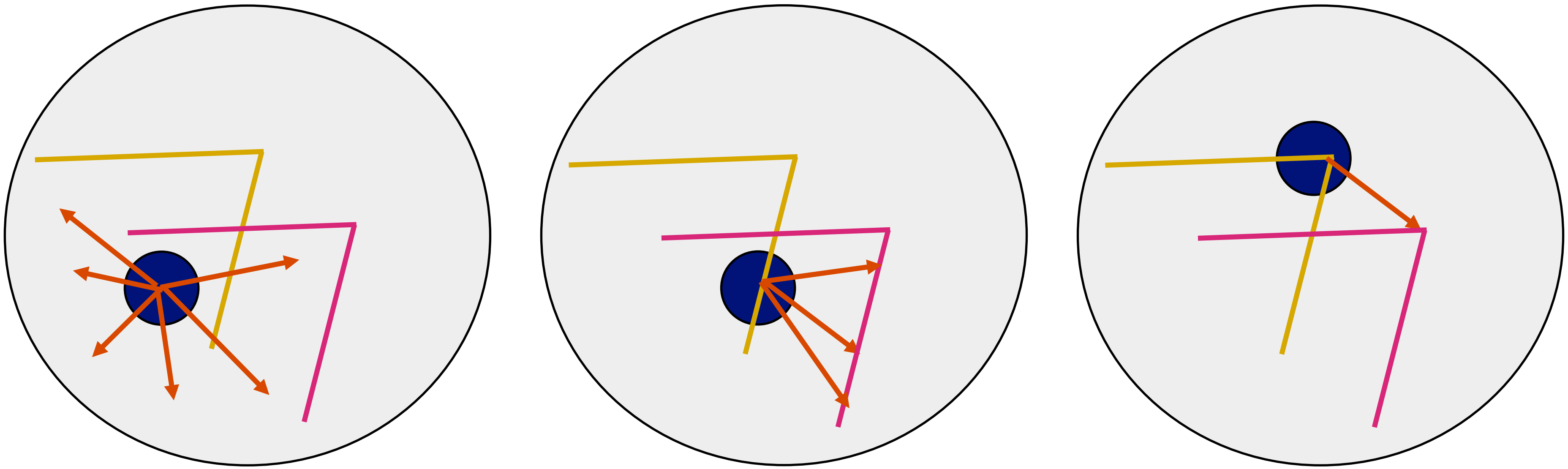
The Aperture Problem



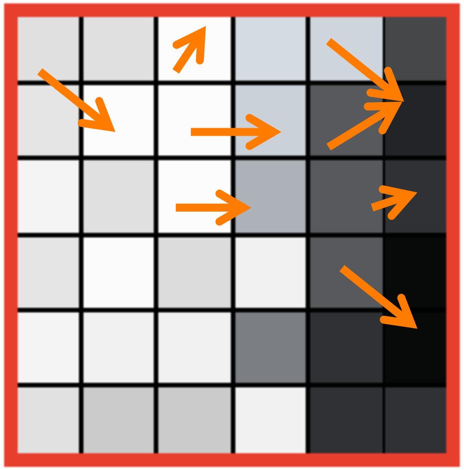
The Aperture Problem



Optical flow uncertaintly



Overcoming the aperture effect

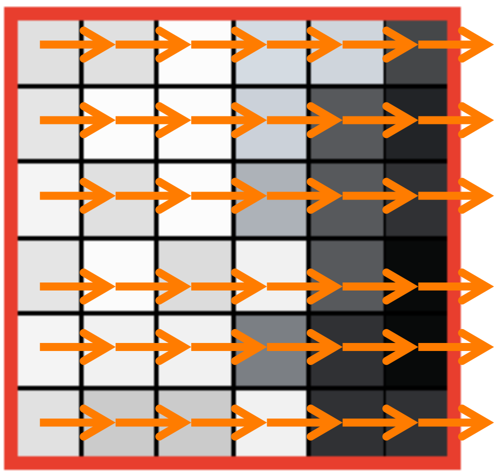


Brightness constancy constraint

$$I_x(p_i)u_i + I_x(p_i)v_i + I_t(p_i) = 0$$

Overcoming the aperture effect

united we move



Brightness constancy constraint

$$I_x(p_i)u + I_x(p_i)v + I_t(p_i) = 0$$

Lucas-Kanade

$$I_x(p_i)u + I_x(p_i)v + I_t(p_i) = 0$$

$$\begin{bmatrix} I_{x,1} & I_{y,1} \\ & \vdots \\ I_{x,25} & I_{y,25} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -I_{t,1} \\ \vdots \\ -I_{t,25} \end{bmatrix}$$

Rewrite: $\mathbf{Au} = \mathbf{b}$ 25 equations, 2 unknows

Residuals: $\epsilon = \mathbf{Au} - \mathbf{b}$

Cost: $\epsilon^T \epsilon = \mathbf{b}^T\mathbf{b} - 2\mathbf{u}^T\mathbf{A}^T\mathbf{b} + \mathbf{u}^T\mathbf{A}^T\mathbf{Au}$

Minimization: $\mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{b}$

B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. IJCAI, 1981.

Lucas-Kanade, continued

$$\mathbf{A}^T\mathbf{A}\mathbf{u} = \mathbf{A}^T\mathbf{b}$$

$$\mathbf{u} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$$

Is it invertible?

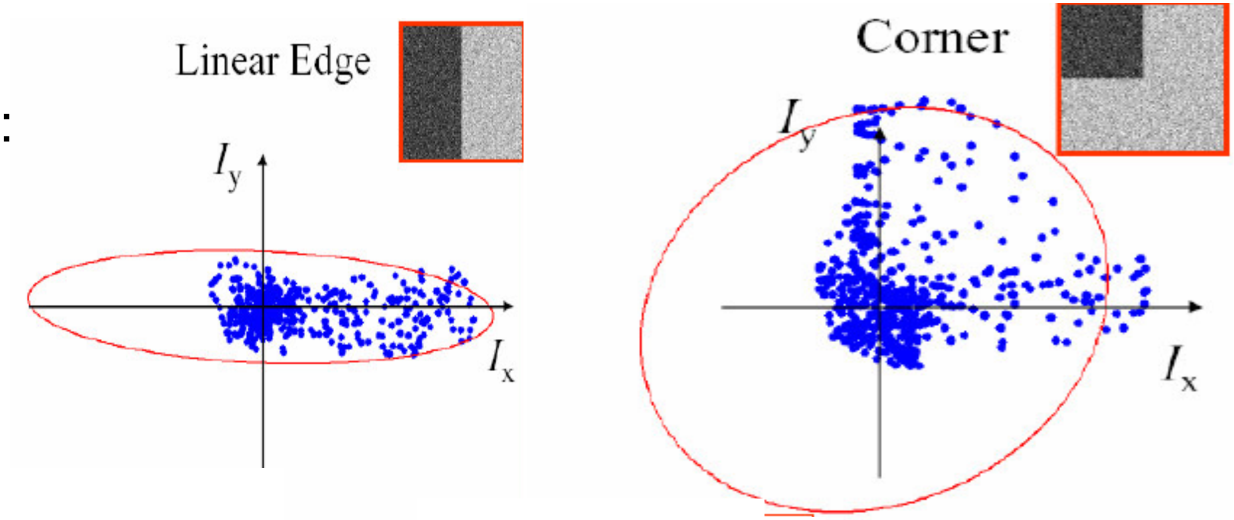
$$\mathbf{A} = \begin{bmatrix} I_{x,1} & I_{y,1} \\ & \vdots & \\ I_{x,25} & I_{x,25} \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} \sum_i I_{x,i}^2 & \sum_i I_{x,i}I_{y,i} \\ \sum_i I_{x,i}I_{y,i} & \sum_i I_{y,i}^2 \end{bmatrix}$$

B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. IJCAI, 1981.

Second Moment Matrix

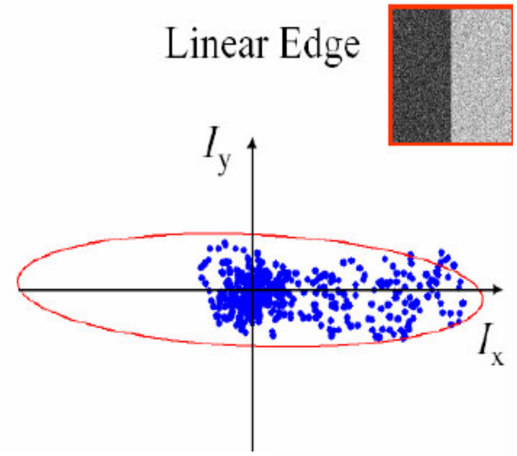
Distribution of gradients:

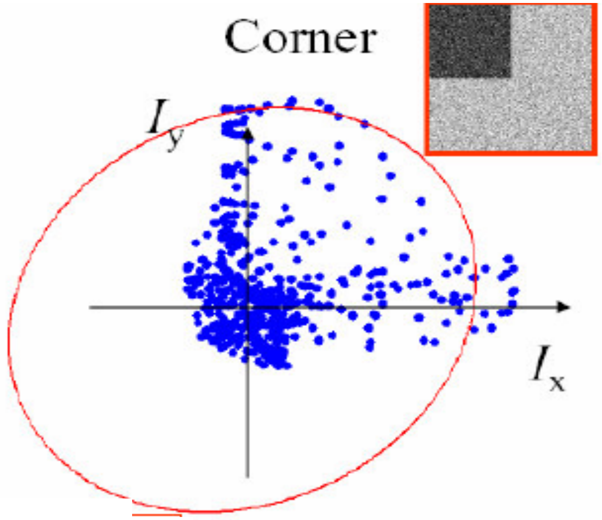


$$J = \begin{bmatrix} \sum_{x',y'} I_x^2 & \sum_{x',y'} I_x I_y \\ \sum_{x',y'} I_x I_y & \sum_{x',y'} I_y^2 \end{bmatrix}$$

Second Moment Matrix

Distribution of gradients:

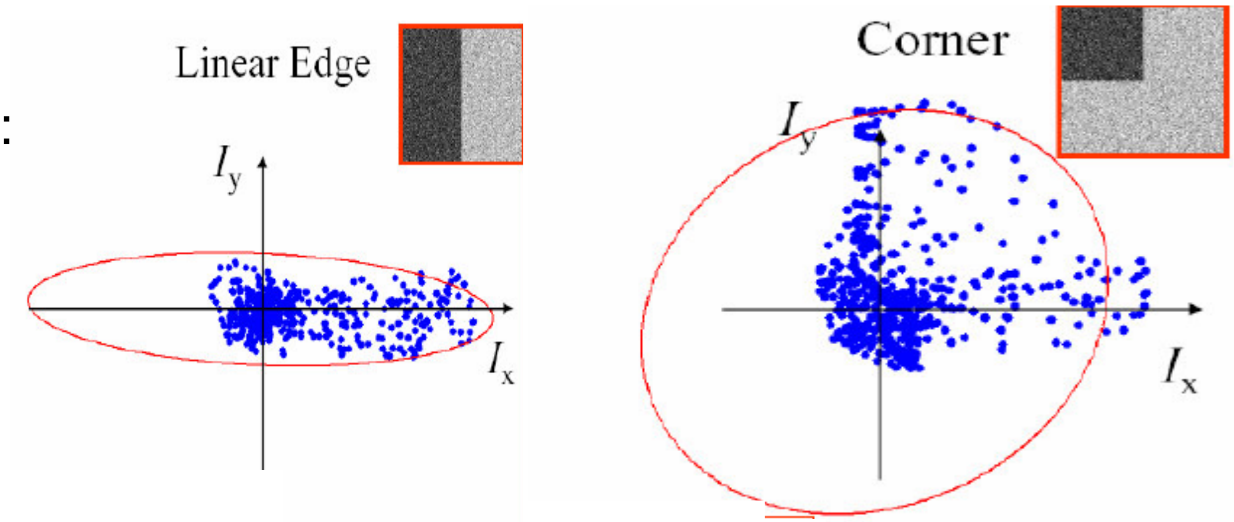




$$J = \sum_{x',y'} (I_x, I_y)^T (I_x, I_y)$$

Second Moment Matrix

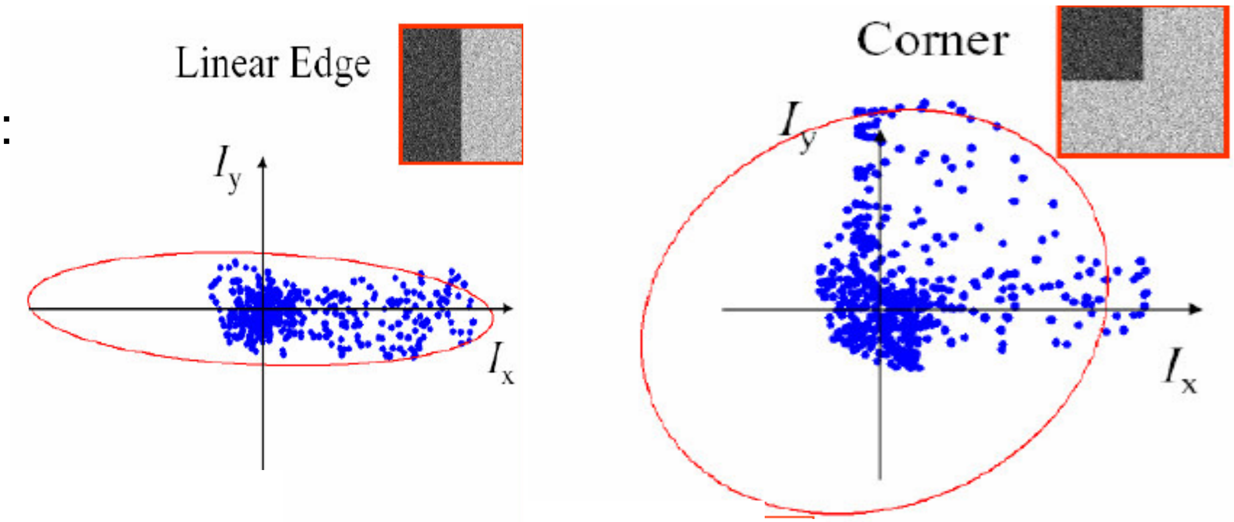
Distribution of gradients:



$$J = \sum_{x',y'} \left(\nabla G_\sigma * u\right)^T \left(\nabla G_\sigma * u\right)$$

Second Moment Matrix

Distribution of gradients:

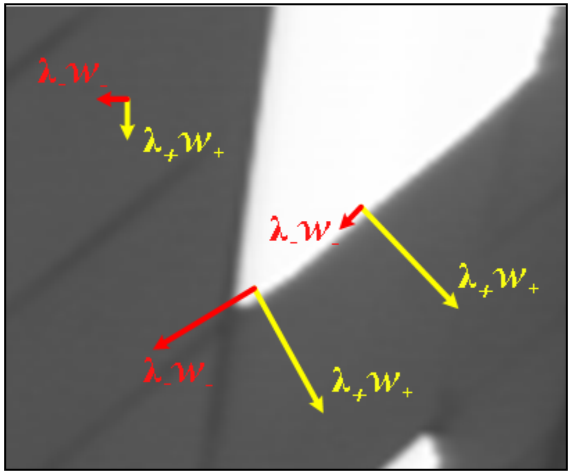


$$J = G_\rho * \left[(\nabla G_\sigma * u)^T (\nabla G_\sigma * u) \right]$$

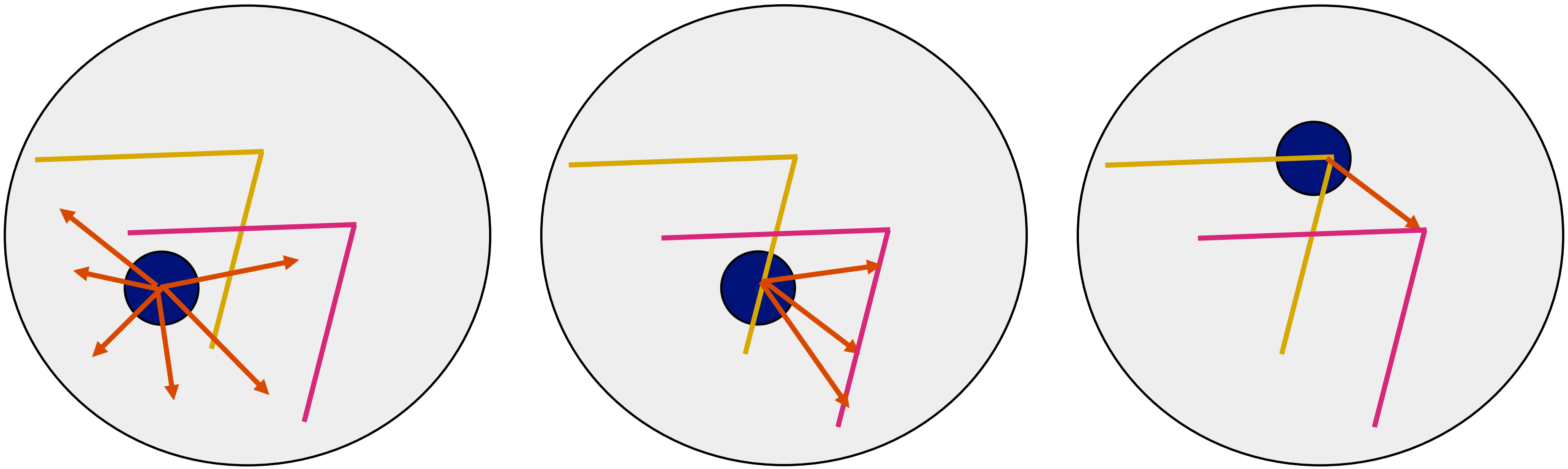
Second Moment Matrix

$$J = G_\rho * \left[(\nabla G_\sigma * u)^T (\nabla G_\sigma * u) \right]$$

- Eigenvectors w_+, w_- : directions of maximal and minimal variation of u
- Eigenvalues: amounts of minimal and maximal variation u



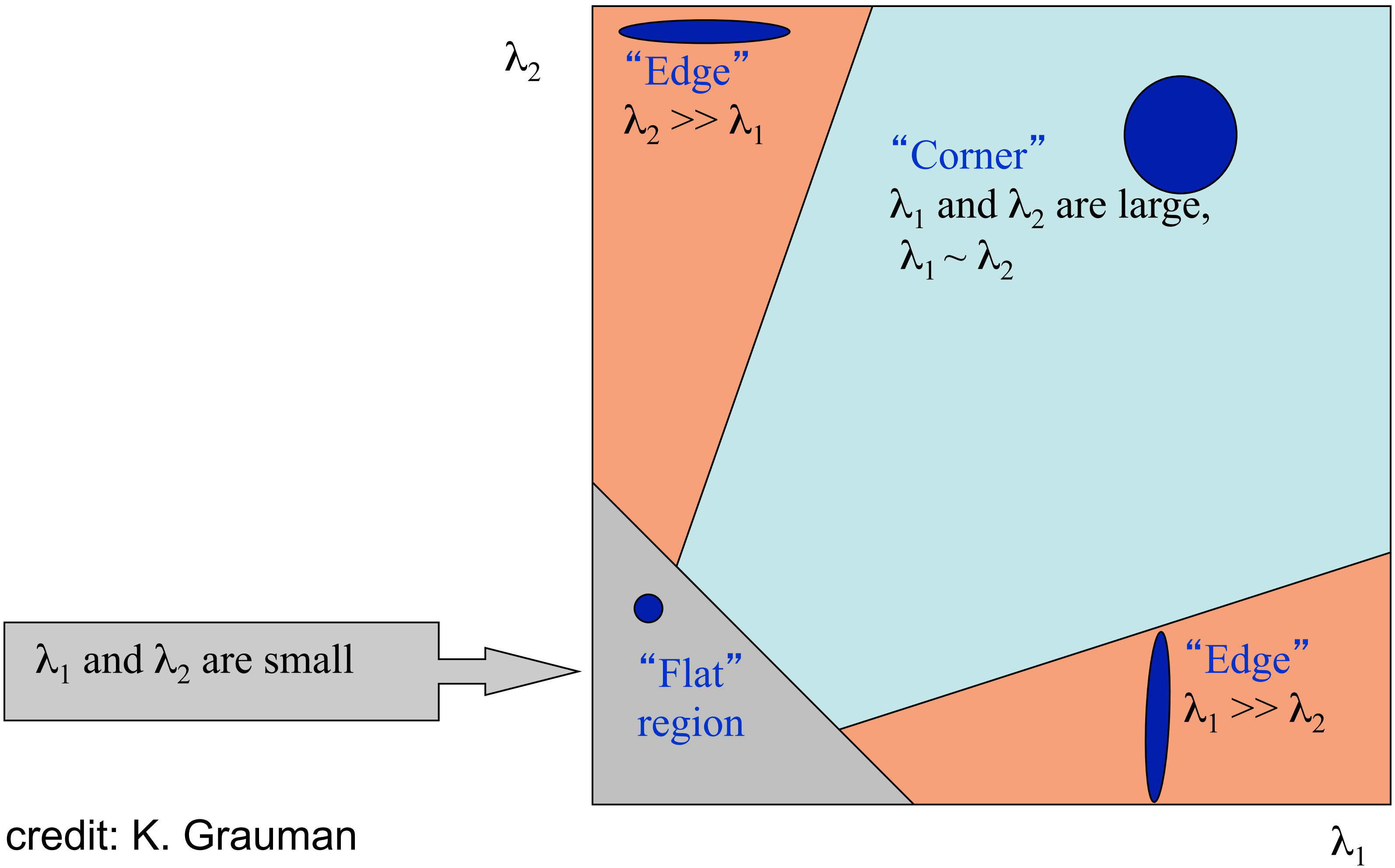
Local structure and motion estimation



$\lambda_1 \simeq 0, \lambda_2 \simeq 0$ λ_1 large, $\lambda_2 \simeq 0$ $\lambda_1, \lambda_2 :$ large

Interpreting the eigenvalues

Classification of image points using eigenvalues of the second moment matrix:

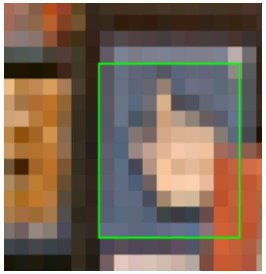


Slide credit: K. Grauman

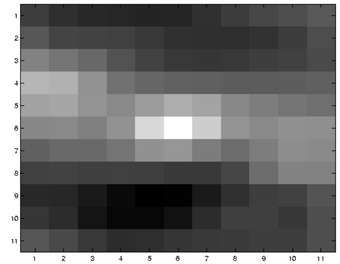
SSD measure

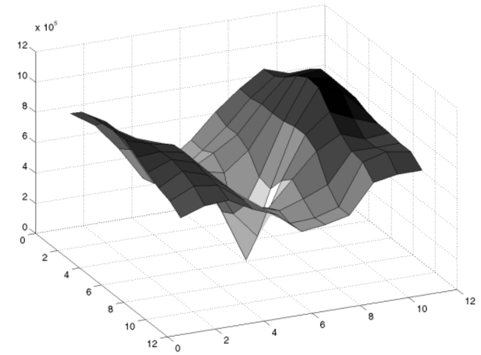
- Sum of squared differences

$$E(x,y;d) = \sum_{(x',y') \in N(x,y)} [I_L(x'+d,y') - I_R(x',y')]^2$$



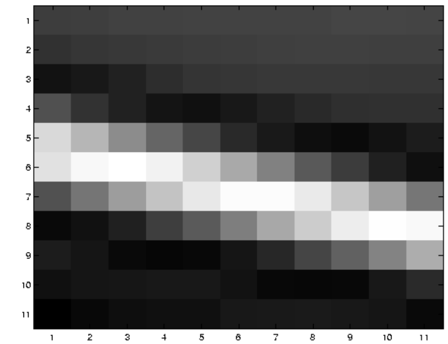
High-Texture Region

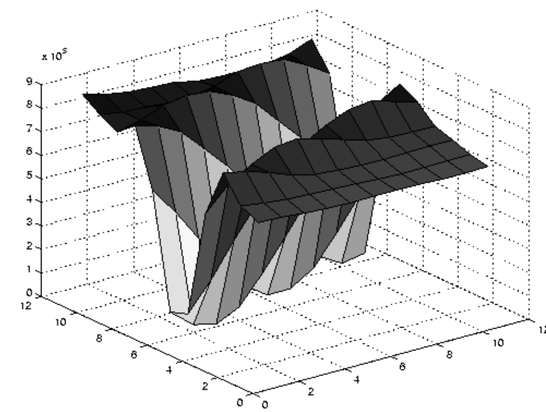




– Gradients are different, large magnitude
– Large λ_1 , large λ_2

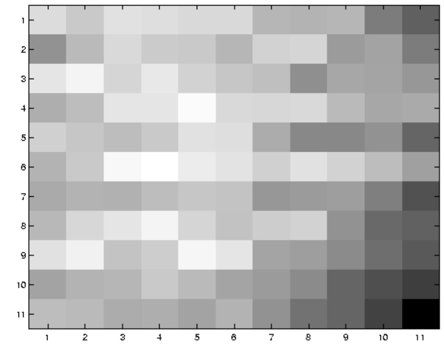
Edge

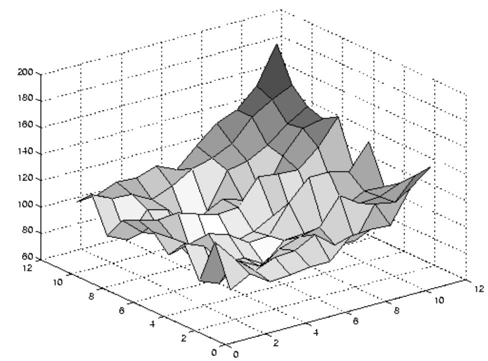




– Gradients very large or very small
– Large λ_1 , small λ_2

Low-Texture Region





– Gradients have small magnitude
– Small λ_1 , small λ_2

Shi-Tomasi feature tracker

1. Find good features (min eigenvalue of 2×2 Hessian)
2. Use Lucas-Kanade to track with pure translation
3. Use affine registration with first feature patch
4. Terminate tracks whose dissimilarity gets too large
5. Start new tracks when needed

Tracking example

Figure 1: Three frame details from Woody Allen's *Manhattan*. The details are from the 1st, 11th, and 21st frames of a subsequence from the movie.

Figure 2: The traffic sign windows from frames 1,6,11,16,21 as tracked (top), and warped by the computed deformation matrices (bottom).

J. Shi and C. Tomasi. [Good Features to Track](). CVPR 1994.

Lucas-Kanade and the aperture effect

Brightness constancy + neighboring pixels have same (u,v)

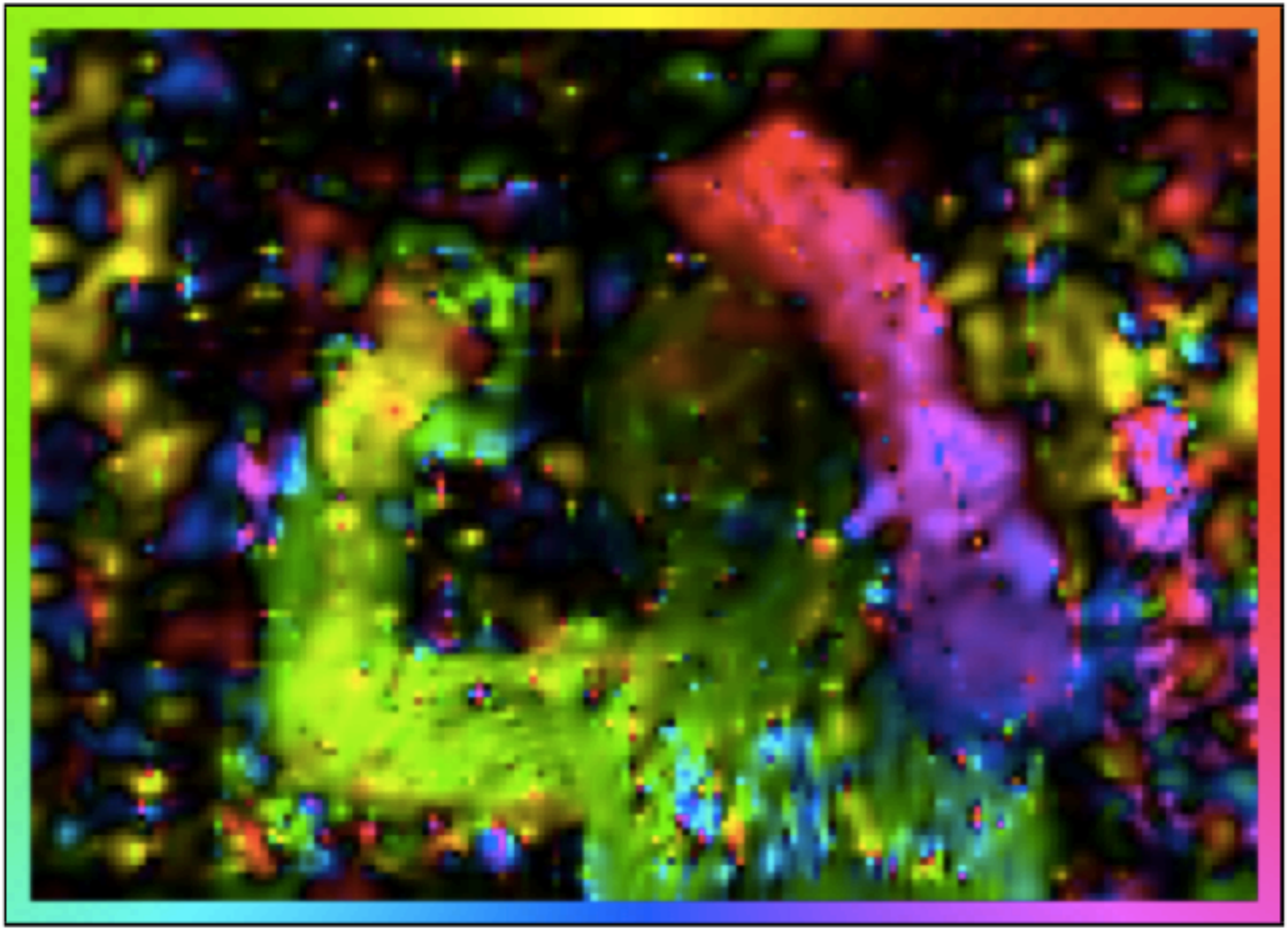
$$I_x(p_i)u + I_x(p_i)v + I_t(p_i) = 0$$

5x5 window: 25 equations, 2 unknowns

$$\begin{bmatrix} I_x(\mathbf{p_1}) & I_y(\mathbf{p_1}) \\ I_x(\mathbf{p_2}) & I_y(\mathbf{p_2}) \\ \vdots & \vdots \\ I_x(\mathbf{p_{25}}) & I_y(\mathbf{p_{25}}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p_1}) \\ I_t(\mathbf{p_2}) \\ \vdots \\ I_t(\mathbf{p_{25}}) \end{bmatrix}$$

Solve the system around each pixel separately

Dense Lukas Kanade



Horn-Schunk and the aperture effect

Brightness constancy + flow smoothness:

$$J(u,v) = \frac{1}{2}\iint_\Omega \left[I_t + \nabla I \cdot (u,v)\right]^2 + \lambda\left[(\nabla u)^2 + (\nabla v)^2\right] \mathrm{d}x\mathrm{d}y$$

Minimize with respect to $u(x,y), v(x,y)$

Euler-Lagrange derivative:

$$J_u = \frac{\partial J}{\partial u} - \frac{\partial}{\partial x}\frac{\partial J}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial J}{\partial u_y}$$

Minimum condition $J_u = 0, \quad J_v = 0$

B.K.P. Horn and B.G. Schunck, "Determining optical flow." Artificial Intelligence, 1981.

Horn-Schunk results

Lucas-Kanade revisited

- General expression for LK criterion:

$$\begin{aligned} E_{LK}(u,v)|_{x,y} &= \sum_{\mathcal{N}(x,y)} (f_x u + f_y v + f_t)^2 \\ &= K_\rho * (f_x u + f_y v + f_t)^2\big|_{x,y} \end{aligned}$$

- LK flow

$$\left[\begin{array}{cc} K_\rho * f_x^2 & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * (f_y)^2 \end{array} \right] \left[\begin{array}{c} u \\ v \end{array} \right] = - \left[\begin{array}{c} K_\rho * (f_x f_t) \\ K_\rho * (f_y f_t) \end{array} \right]$$

- How does this compare with Horn-Schunk criterion?

Lucas-Kanade meets Horn-Schunk

Introduce: $$w \doteq \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \qquad \nabla_3 f \doteq \begin{bmatrix} f_x \\ f_y \\ f_t \end{bmatrix}$$

$$|\nabla w|^2 \doteq |\nabla u|^2 + |\nabla v|^2 \qquad J_\rho \doteq G_\rho * \left[\nabla_3 f \nabla_3 f^T\right]$$

Lucas-Kanade

$$E_{LK}(w) = \int_\Omega w^T J_\rho w \; \mathrm{d}x\mathrm{d}y$$

Horn-Schunk

$$E_{HS}(w) = \int_\Omega w^T J_0 w + a|\nabla w| \mathrm{d}x\mathrm{d}y$$

Bruhn-Weickert-Schnoerr

$$E_{BWS}(w) = \int_\Omega w^T J_\rho w + a|\nabla w| \mathrm{d}x\mathrm{d}y$$

A. Bruhn, J. Weickert, C. Schnörr: 'Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods.' IJCV 2005

Horn-Schunk

Bruhn-Weickert-Schnoerr

Bruhn-Weickert-Schnoerr, continued

$$E_{BWS}(w) = \int_\Omega w^T J_\rho w + a|\nabla w| \mathrm{d}x\mathrm{d}y$$

Spatio-temporal regularization

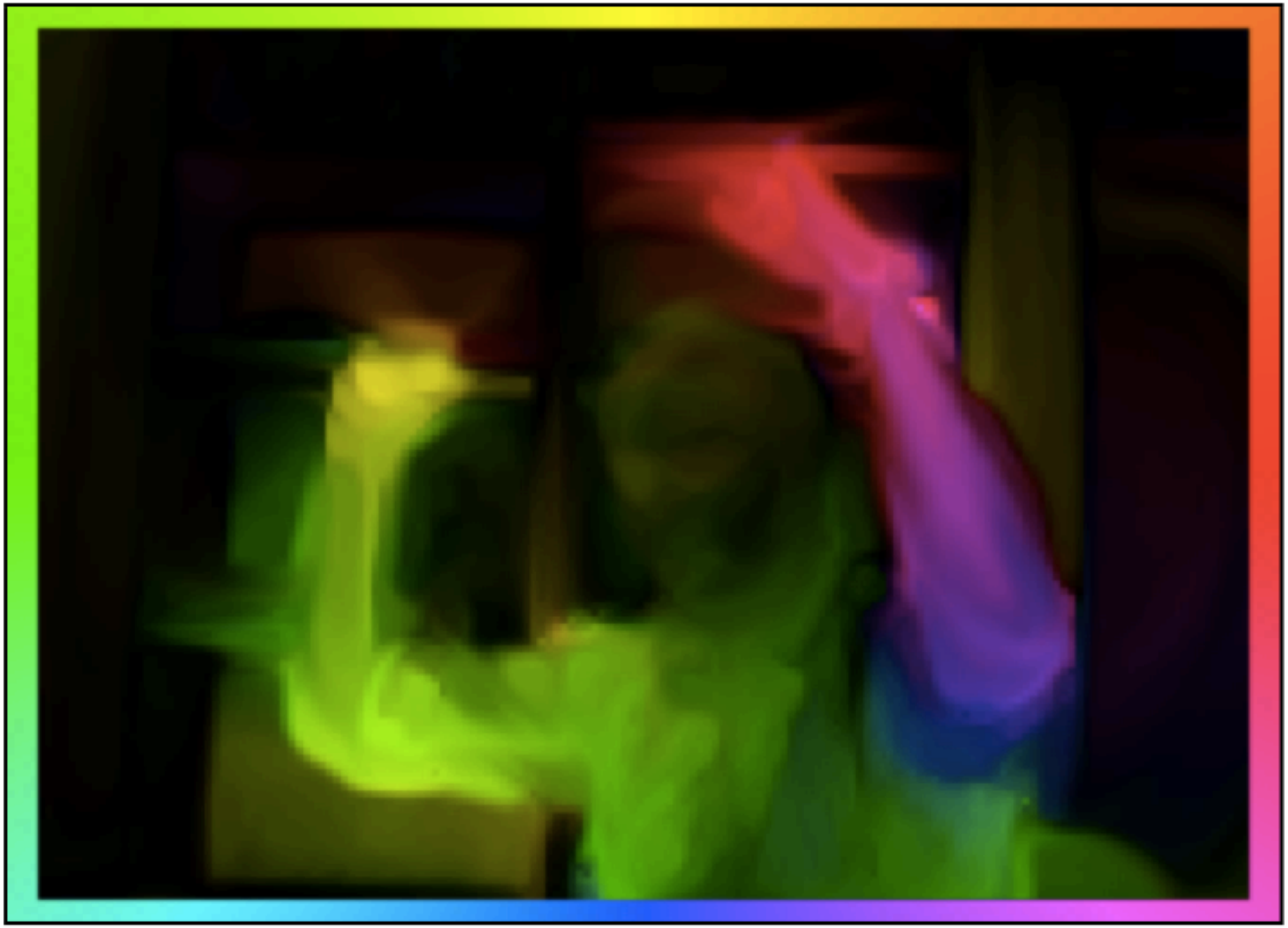
$$E'_{BWS}(w) = \int_{\Omega\times[0,T]} w^T J_\rho w + a|\nabla_3 w| \mathrm{d}x\mathrm{d}y$$

Robust norms:

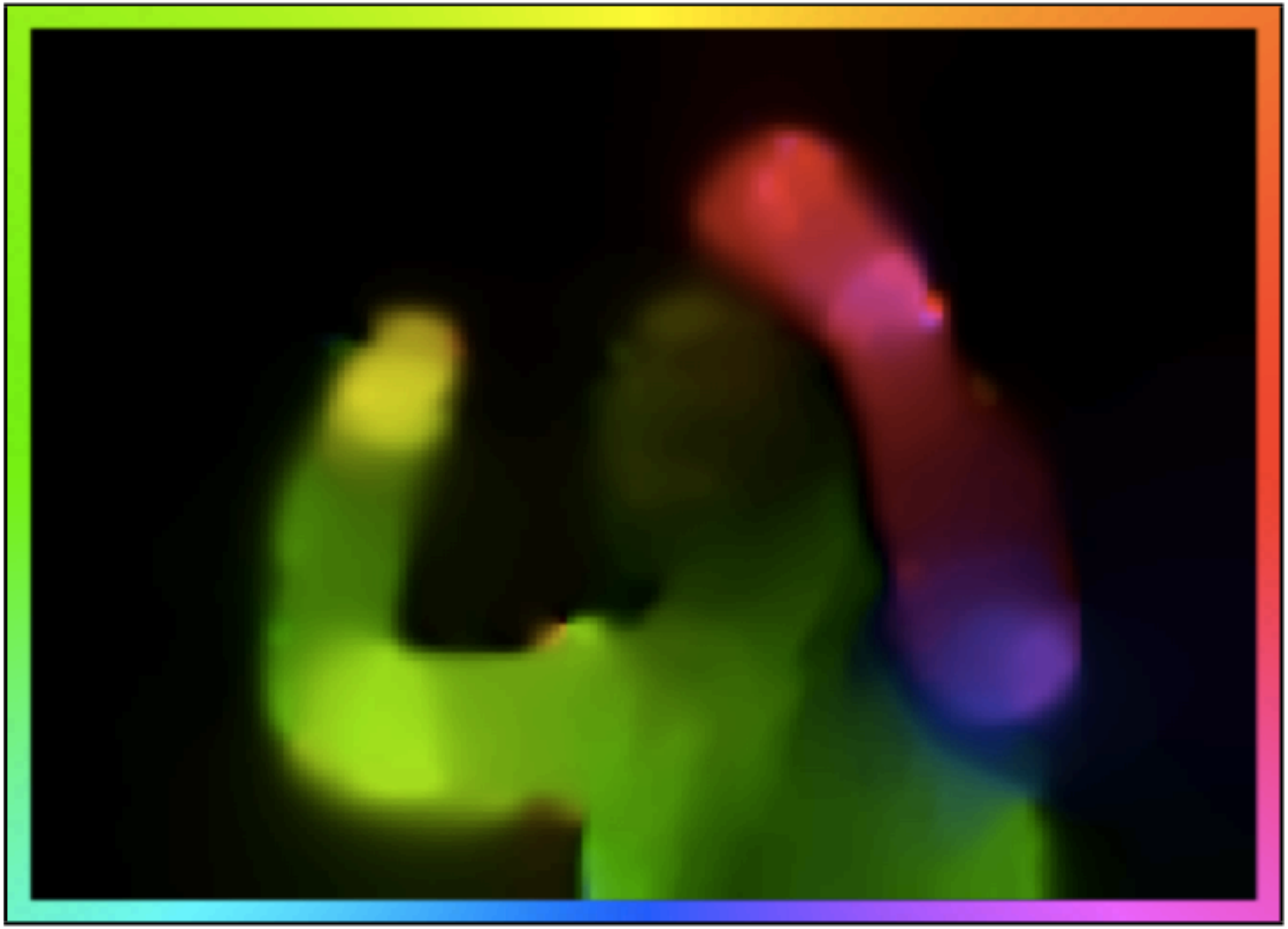
$$E''_{BWS}(w) = \int_{\Omega\times[0,T]} \psi_1\left(w^T J_\rho w\right) + a\psi_2\left(|\nabla_3 w|\right) \mathrm{d}x\mathrm{d}y$$

Bruhn-Weickert-Schnoerr

Bruhn-Weickert-Schnoerr - anisotropic



Bruhn-Weickert-Schnoerr – flow regularization



Numerical solutions

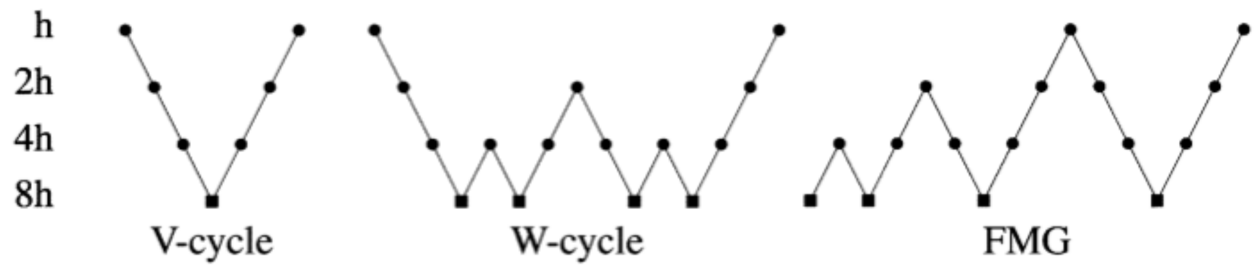
- Euler-Lagrange:

$$I_x^2 u + I_x I_y v = \lambda \nabla^2 u - I_x I_t$$
$$I_x I_y u + I_y^2 v = \lambda \nabla^2 v - I_y I_t$$

- Numerical approximation to Laplacian $\nabla^2 u \simeq 4\hat{u}(x,y) - u(x,y)$

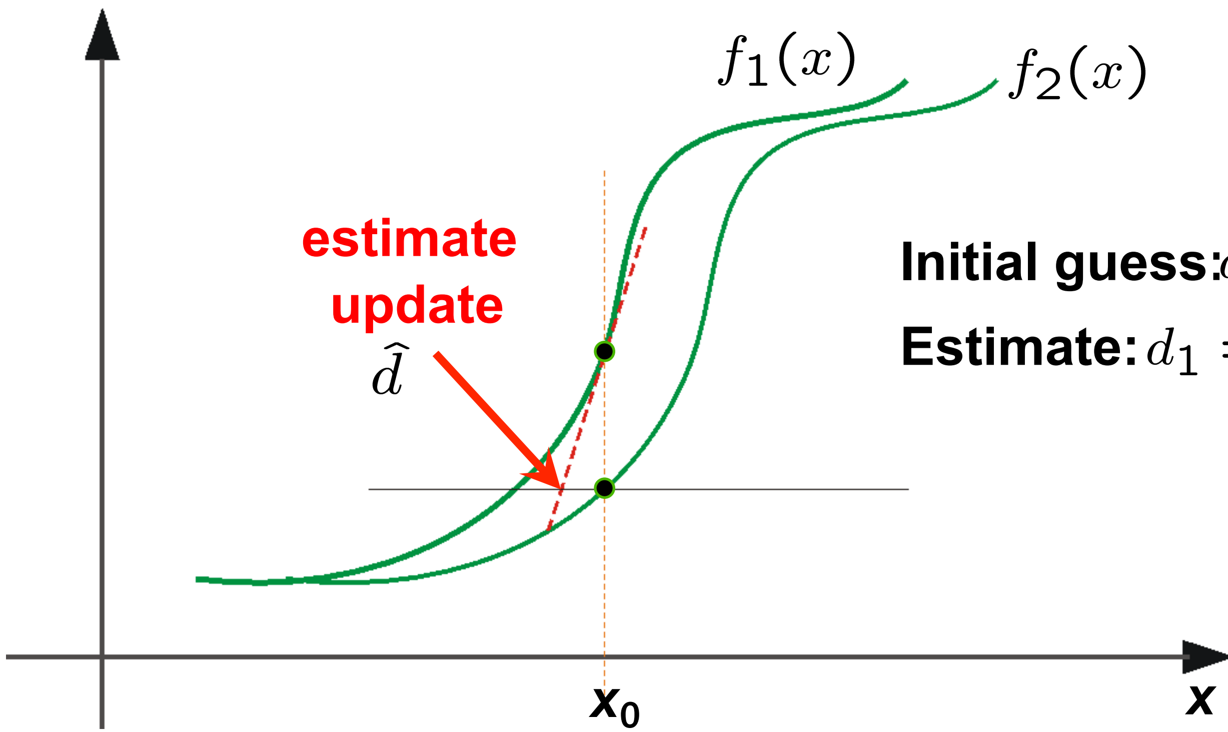
$$\hat{u}(x,y) = \frac{1}{4}\left(u(x-1,y) + u(x+1,y) + u(x,y-1) + u(x,y+1)\right)$$

- Sparse linear system in u,v
  - Gauss-Seidel, Successive Over Relaxation (SOR)
  - Multigrid
  - 12 fps, 2006



A. Brandt, “Multi-level adaptive solutions to boundary-value problems,” Math. Comput., 1977.
D. Terzopoulos, Image Analysis Using Multigrid Relaxation Methods, PAMI, 1986
A. Kenigsberg, R. Kimmel, and I. Yavneh, “A Multigrid Approach for Fast Geodesic Active Contours,” 2001
G. Papandreou and P. Maragos, "Multigrid Geometric Active Contour Models", TIP 2007
A. Bruhn, J. Weickert, T. Kohlberger, C. Schnörr: A multigrid platform for real-time motion computation with discontinuity-preserving variational methods. IJCV 2006

Optical Flow: Iterative Estimation

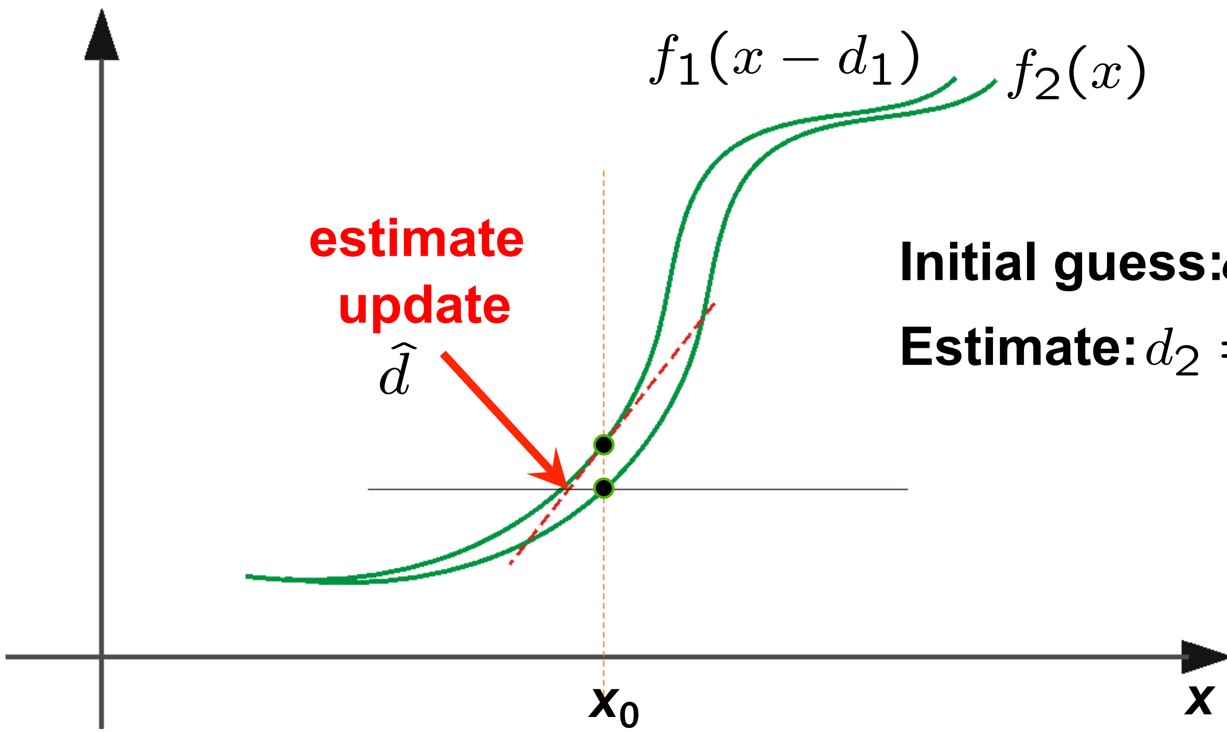


Initial guess: $d_0 = 0$

Estimate: $d_1 = d_0 + \hat{d}$

(using *d* for *displacement* here instead of *u*)

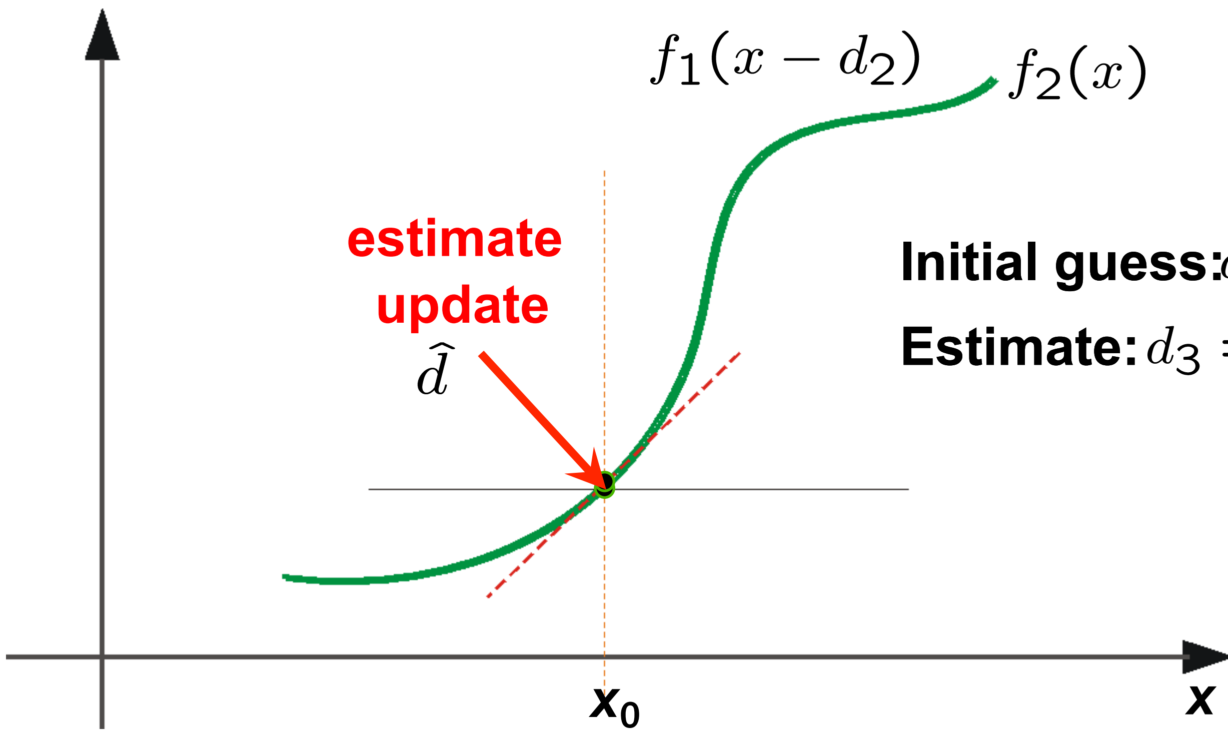
Optical Flow: Iterative Estimation



Initial guess: d_1

Estimate: $d_2 = d_1 + \hat{d}$

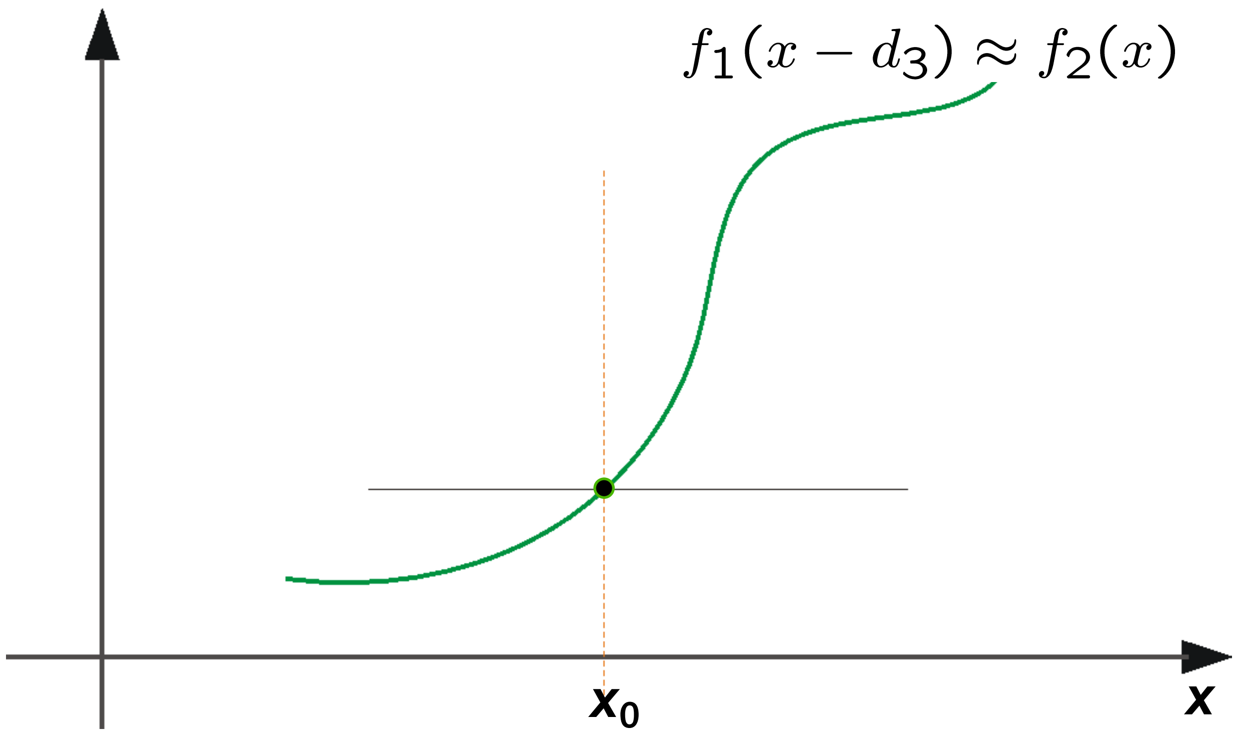
Optical Flow: Iterative Estimation



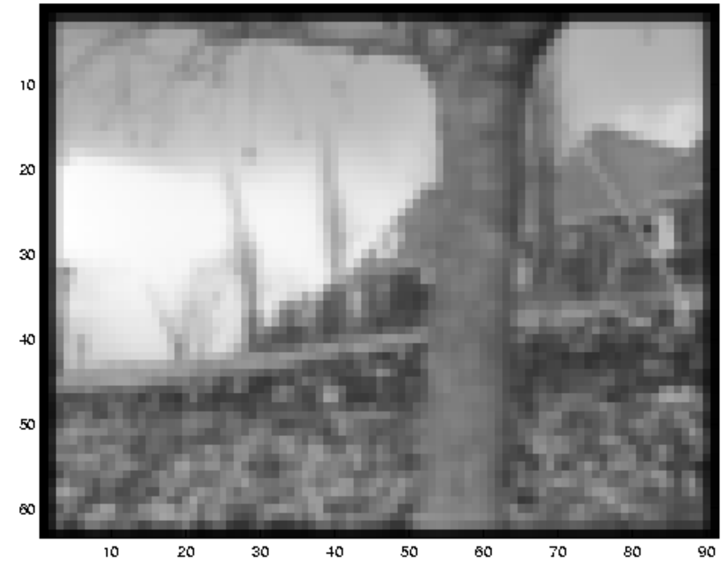
Initial guess: d_2

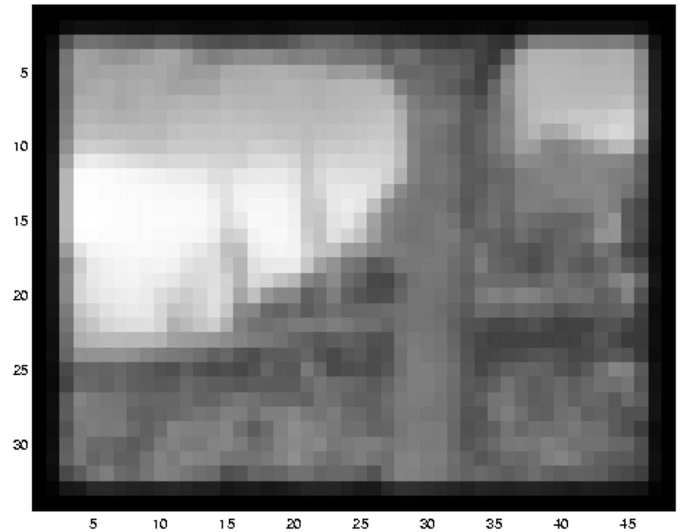
Estimate: $d_3 = d_2 + \hat{d}$

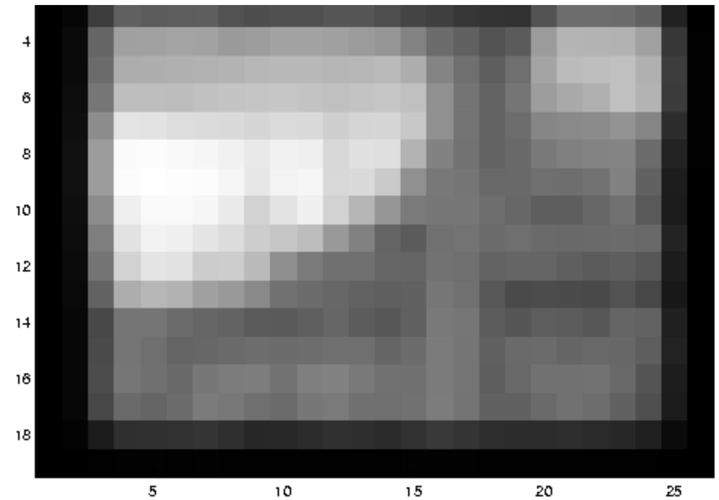
Optical Flow: Iterative Estimation



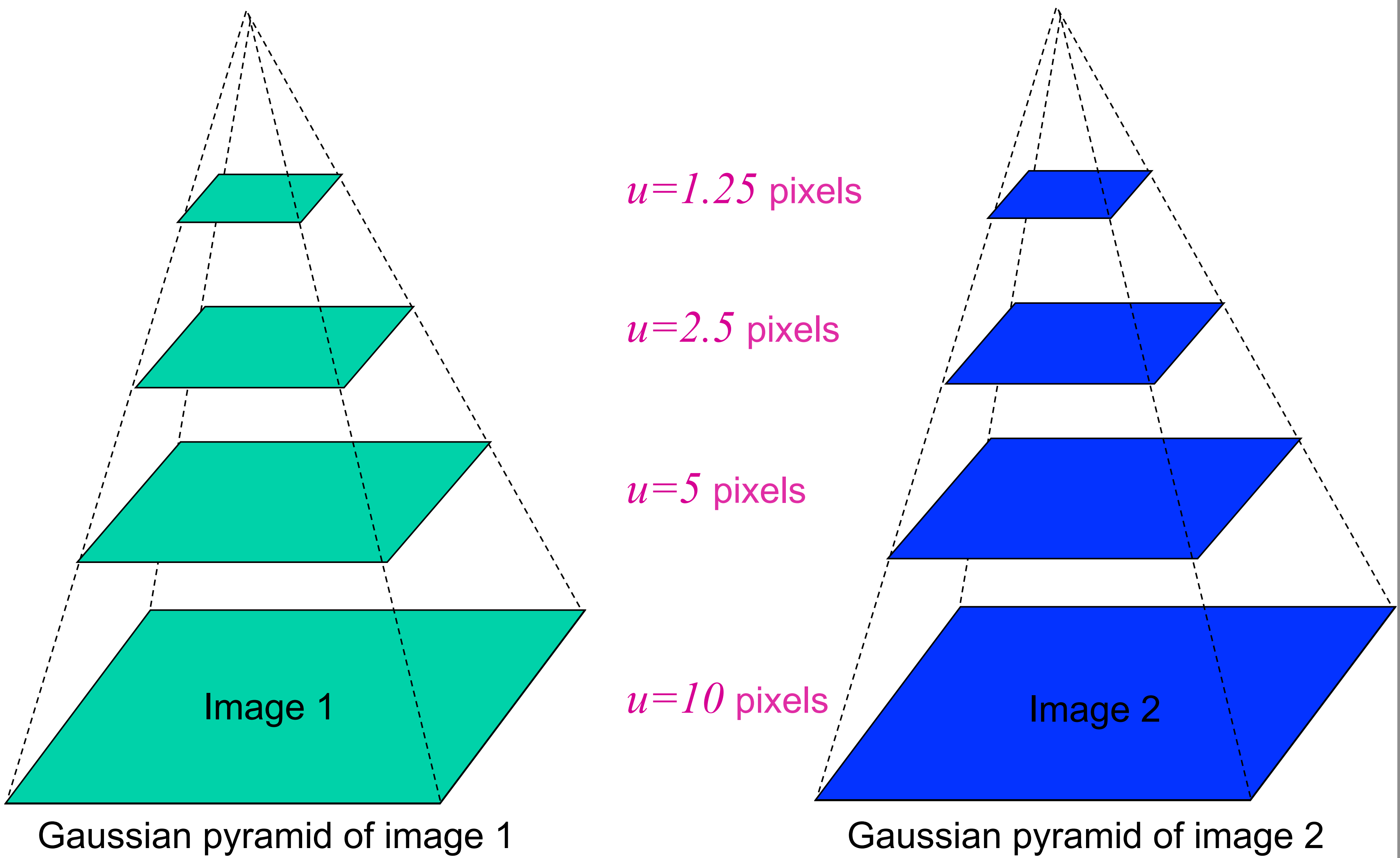
Large Displacements: Reduce Resolution!



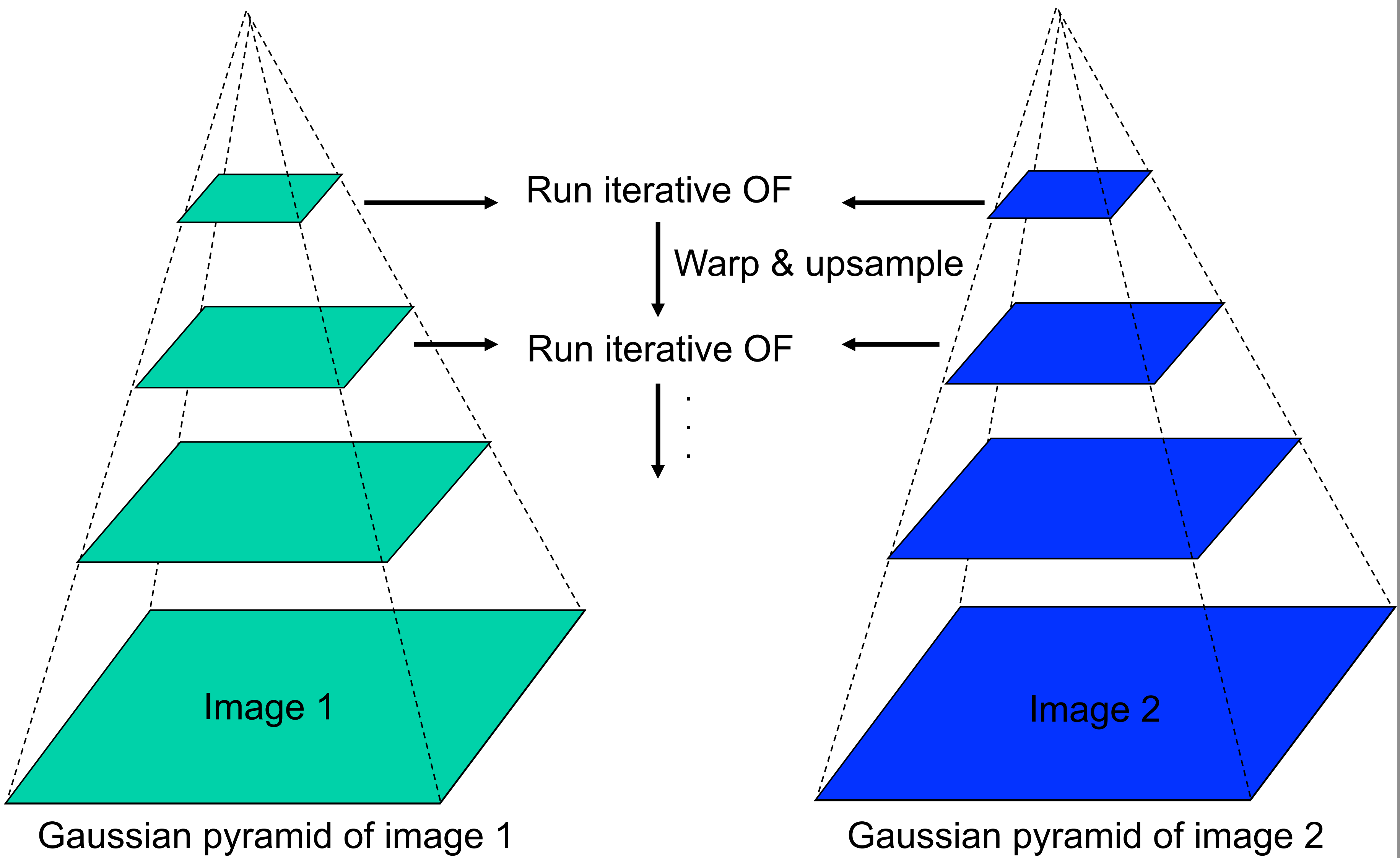




Coarse-to-fine Optical Flow Estimation



Coarse-to-fine Optical Flow Estimation



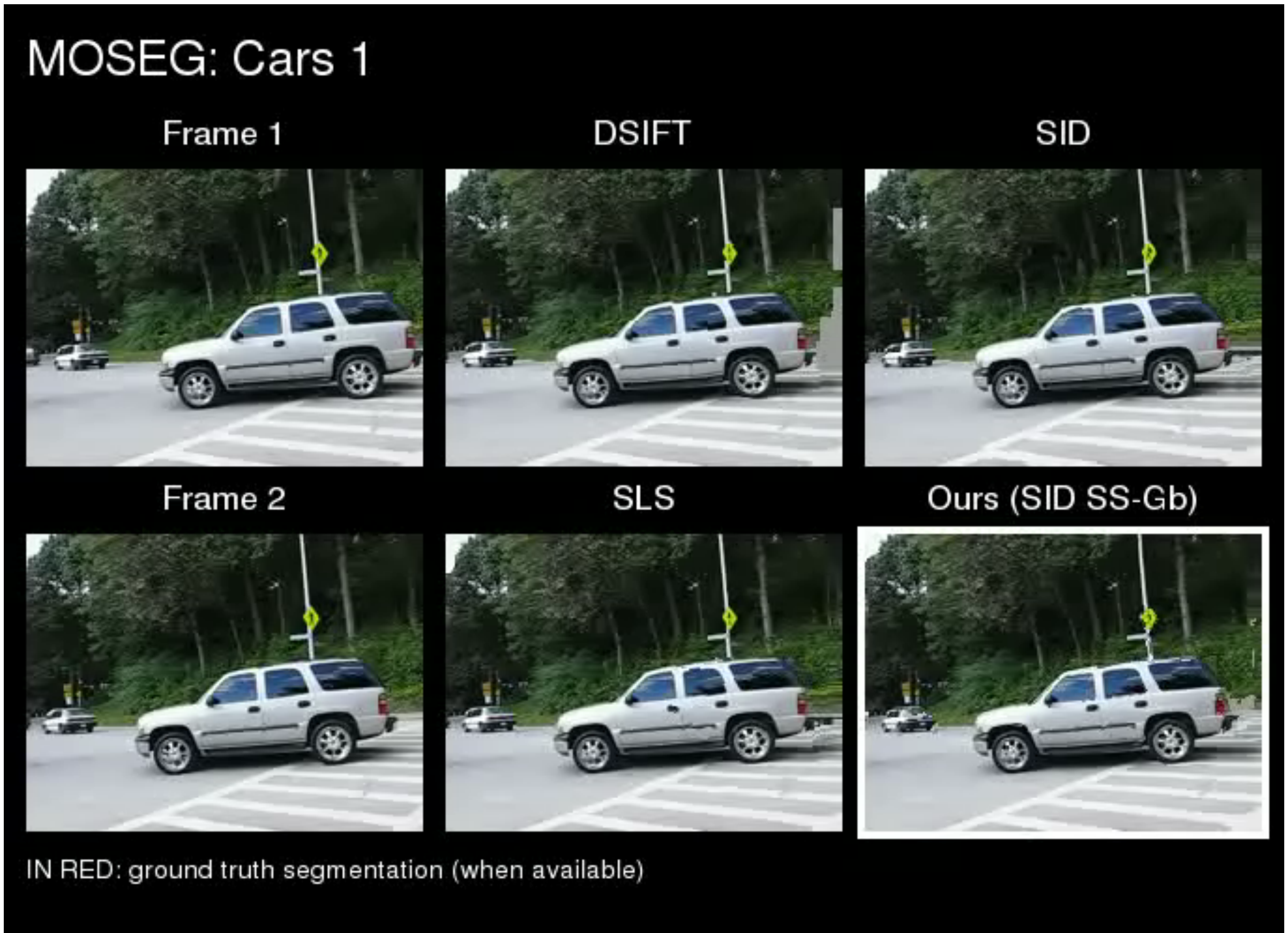
Beyond the brightness constancy constraint

SIFT flow: dense correspondence across different scenes, C Liu, J Yuen, A Torralba, J. Sivic, W. Freeman, ECCV 2008

T. Brox, J. Malik, Large displacement optical flow: descriptor matching in variational motion estimation, PAMI 2011

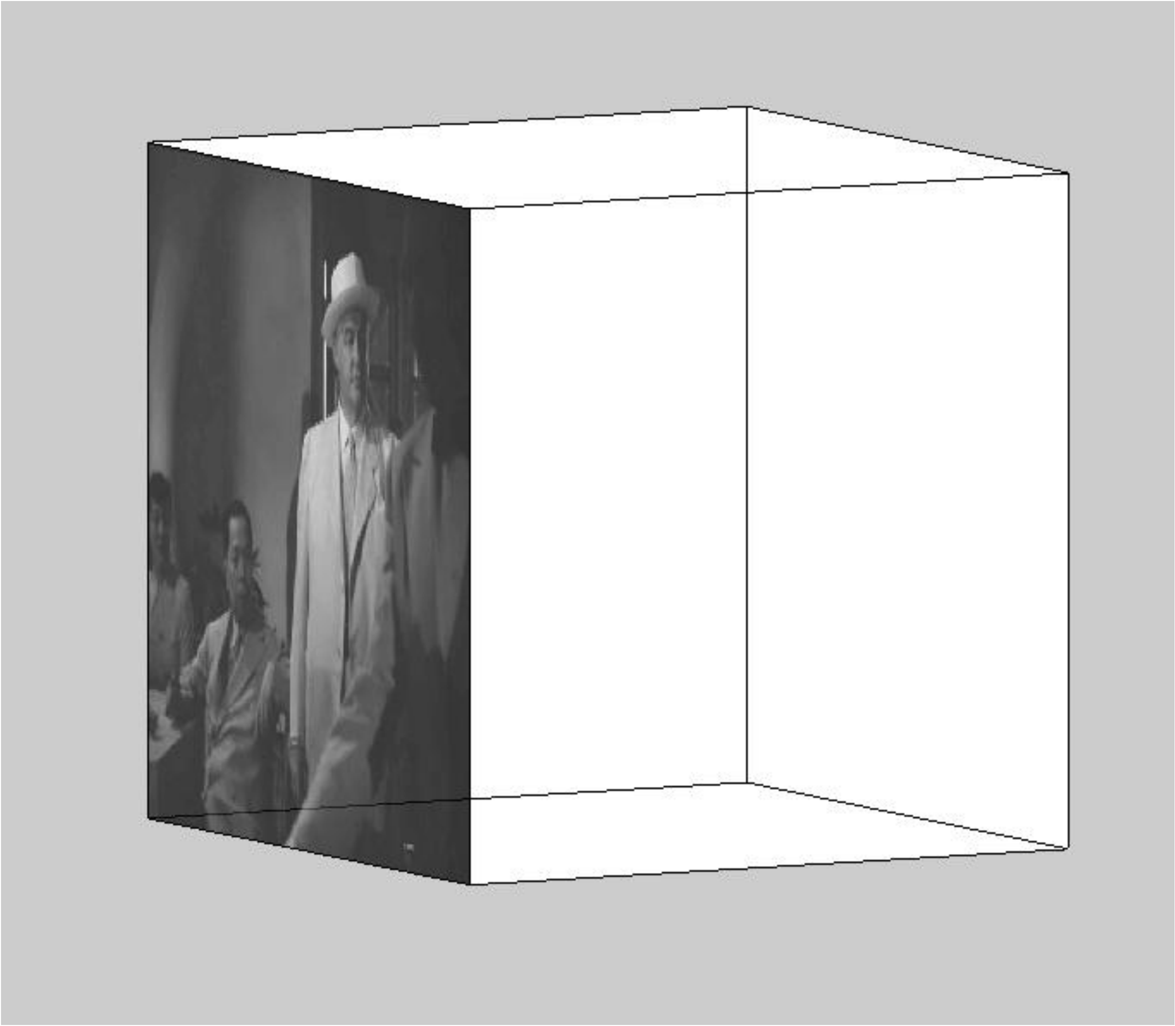
Beyond the brightness constancy constraint

E. Trulls, I. Kokkinos, A. Sanfeliu, and F. Moreno, Dense Segmentation-Aware Descriptors, CVPR 2013



Motion-based action recognition

Discovering Discriminative Action Parts from Mid-Level Video Representations,
M. Raptis, I. Kokkinos and S. Soatto. CVPR, 2012.

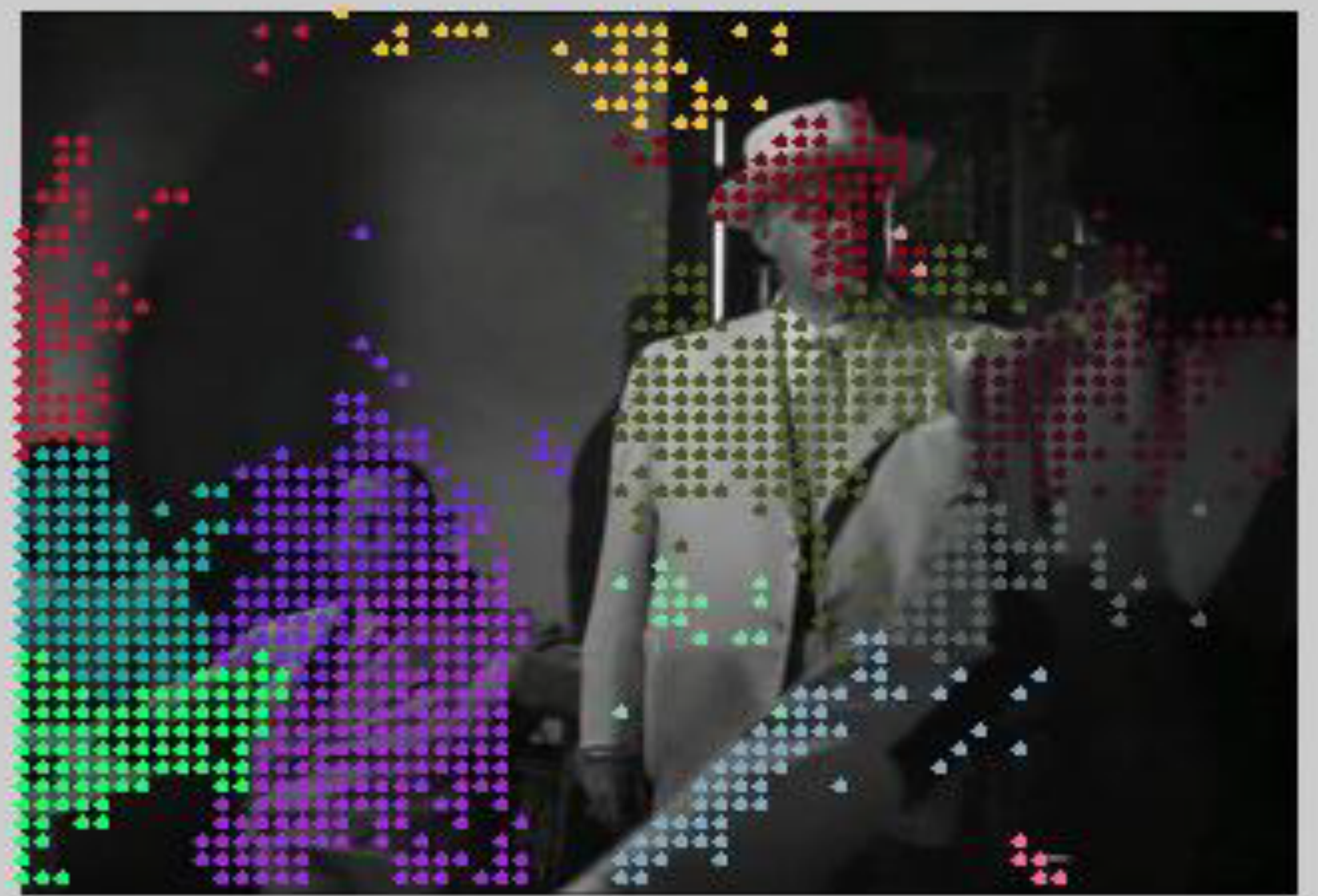


Motion-based action recognition

Discovering Discriminative Action Parts from Mid-Level Video Representations, M. Raptis, I. Kokkinos and S. Soatto. CVPR 12.

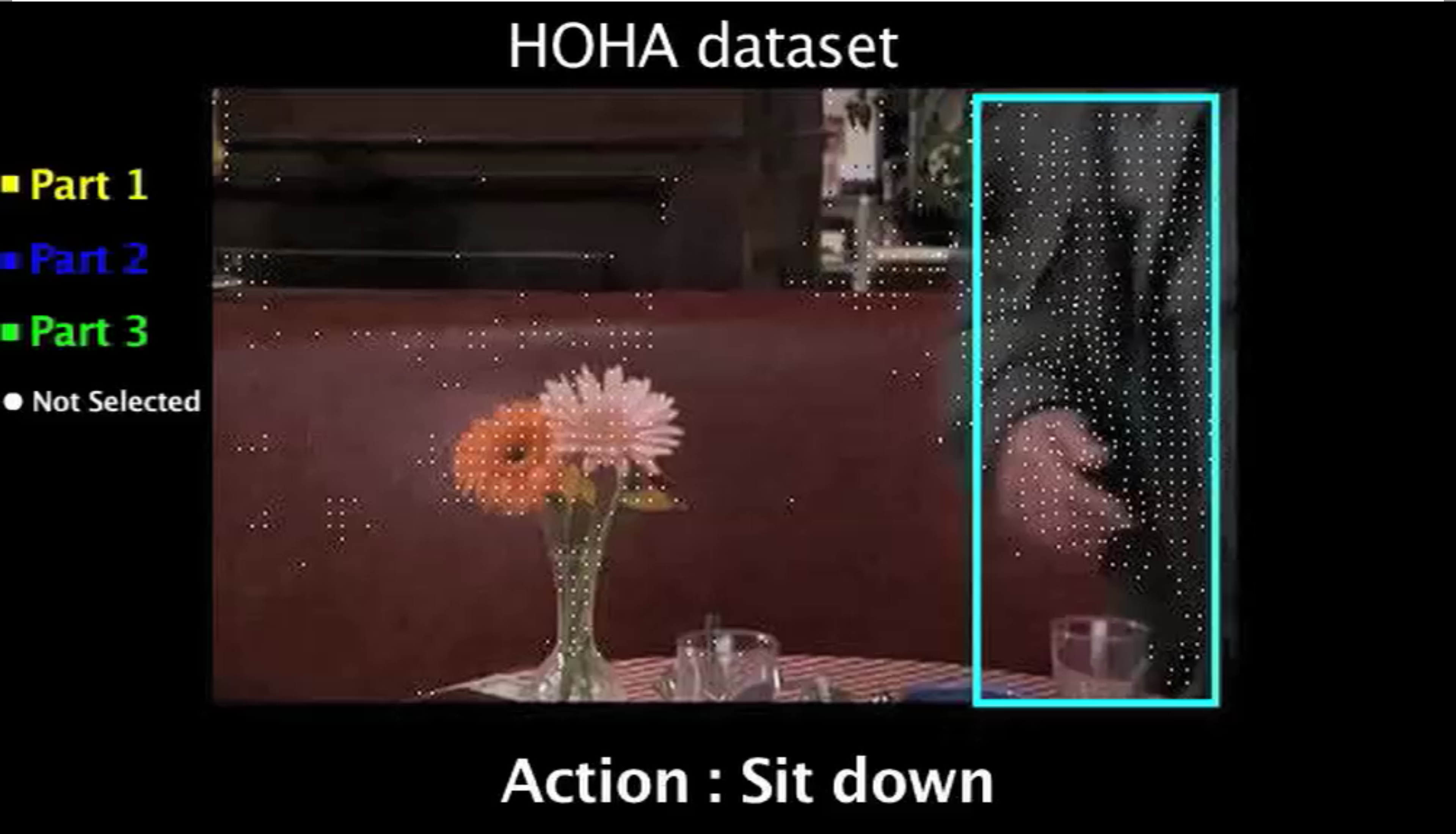
Motion-based action recognition

Discovering Discriminative Action Parts from Mid-Level Video Representations,
M. Raptis, I. Kokkinos and S. Soatto. CVPR, 2012.

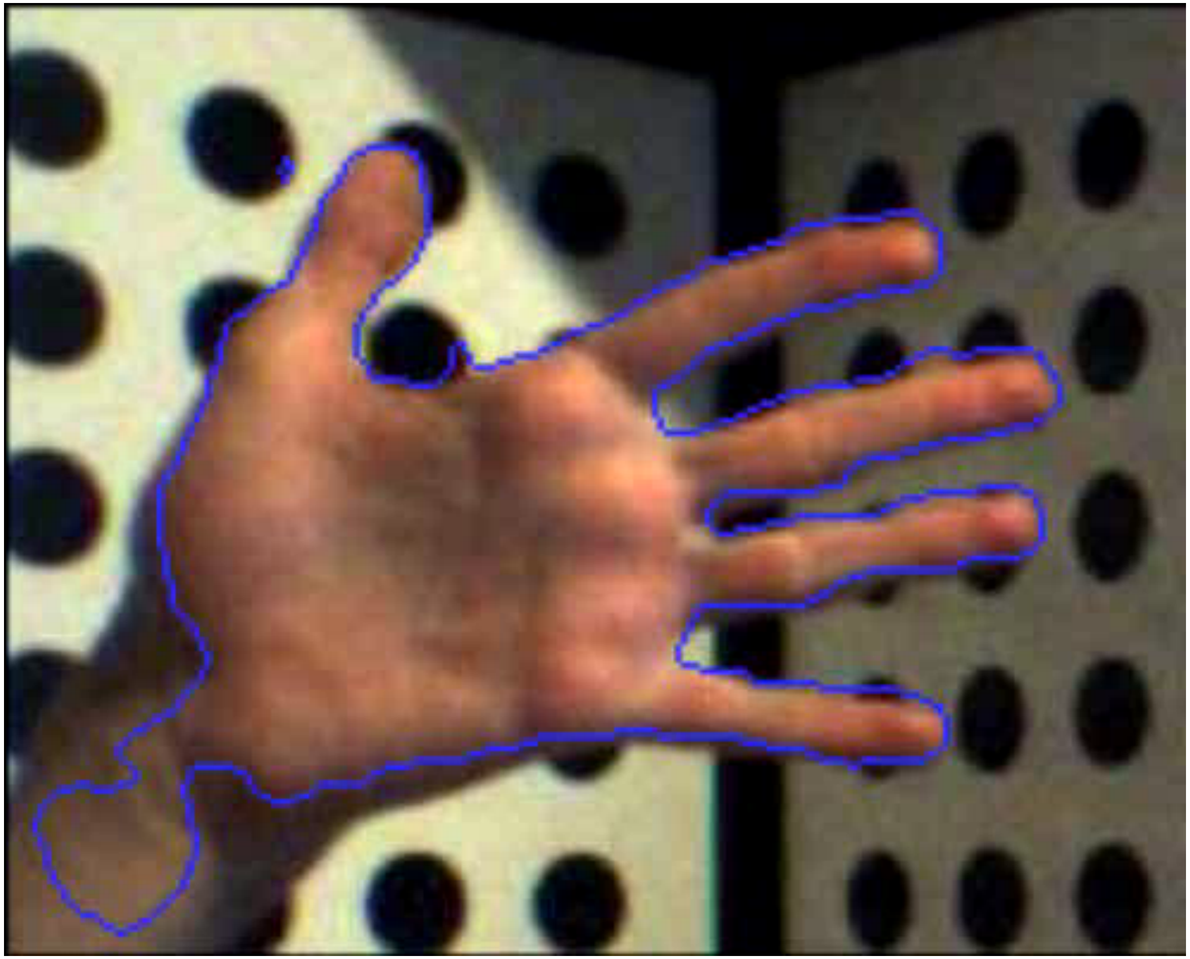


Motion-based action recognition

Discovering Discriminative Action Parts from Mid-Level Video Representations,
M. Raptis, I. Kokkinos and S. Soatto. CVPR, 2012.



Motion-based segmentation



Unsupervised segmentation incorporating colour, texture, and motion. T Brox, M Rousson, R Deriche, J Weickert, IVC 2010

Further Literature Pointers

Layered motion, parametric models, motion segmentation:

Layered Representation for Motion Analysis. CVPR 1993. J. Wang and E. Adelson

Black, M. J. and Anandan, P., The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields, CVIU 1996

D. Cremers, S. Soatto, Motion Competition: A Variational Approach to Piecewise Parametric. Motion Segmentation, IJCV 2004

Unsupervised segmentation incorporating colour, texture, and motion. T Brox, M Rousson, R Deriche, J Weickert, IVC 2010

(wait for R. Szeliski's talk)

Learning:

Roth, S., Black, M.J.: On the spatial statistics of optical flow. IJCV 74, 33–50 (2007)

Discrete optimization: (W. Freeman's talk today)

Fusion Moves for Markov Random Field Optimization. V. Lempitsky, C. Rother, S. Roth, and A. Blake, PAMI 2010.

B. Glocker, N. Komodakis, G. Tziritas, N. Navab, N. Paragios Dense Image Registration through MRFs and Efficient Linear Programming MIA, 2008

Code

- Secrets of optical flow estimation and their principles, Sun., Roth., and Black, CVPR 10
- T. Brox, J. Malik, Large displacement optical flow: descriptor matching in variational motion estimation, PAMI 2011
- Action recognition & descriptor works:
  - http://vision.mas.ecp.fr/Personnel/iasonas/code.html
- Registration service: http://cvn.ecp.fr/
- www.ipol.im