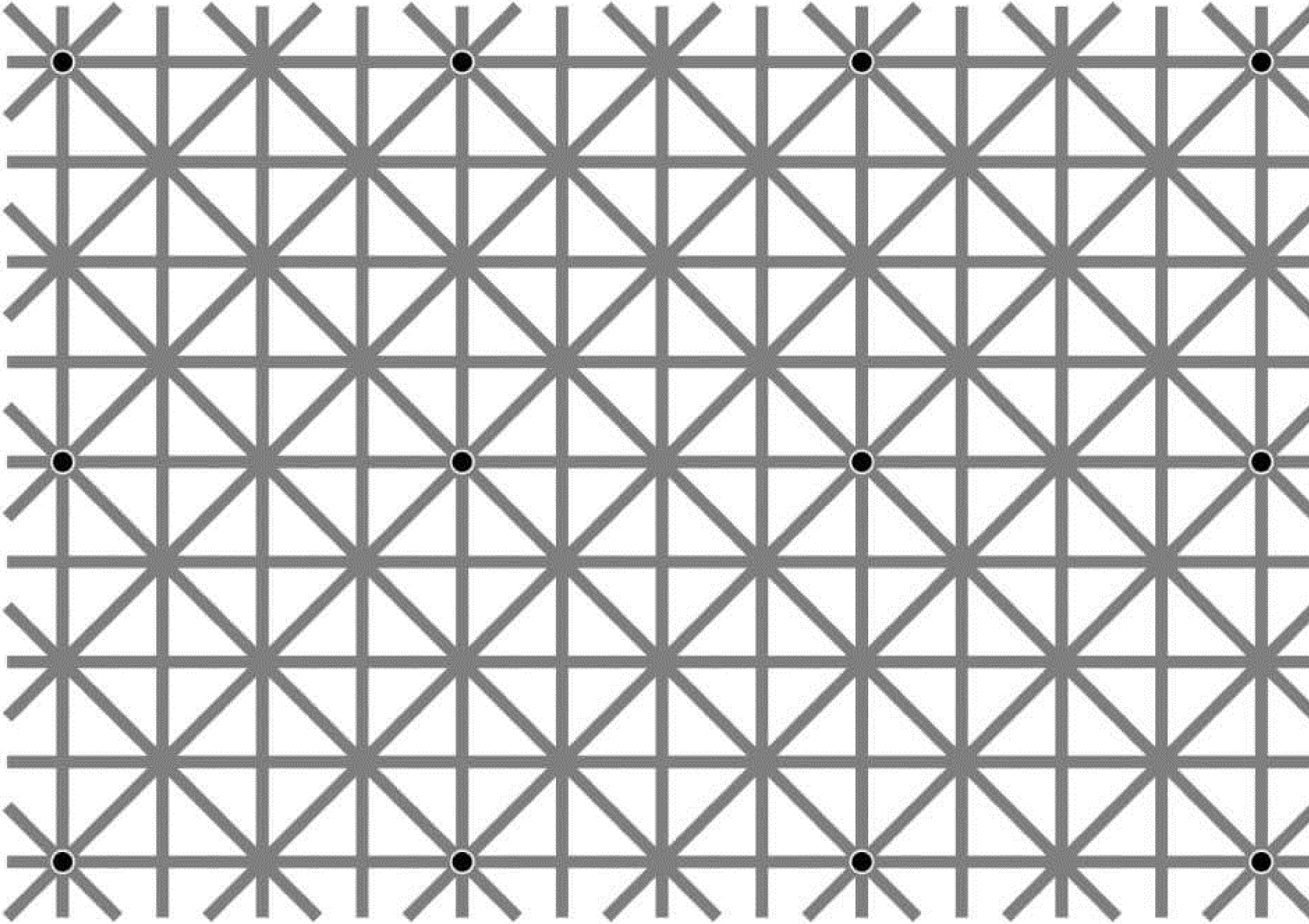


Ninio, J. and Stevens, K. A. (2000) Variations on the Hermann grid: an extinction illusion. Perception, 29, 1209-1217.

Perception, 2000, volume 29, pages 1209 – 1217

DOI:10.1068/p2985

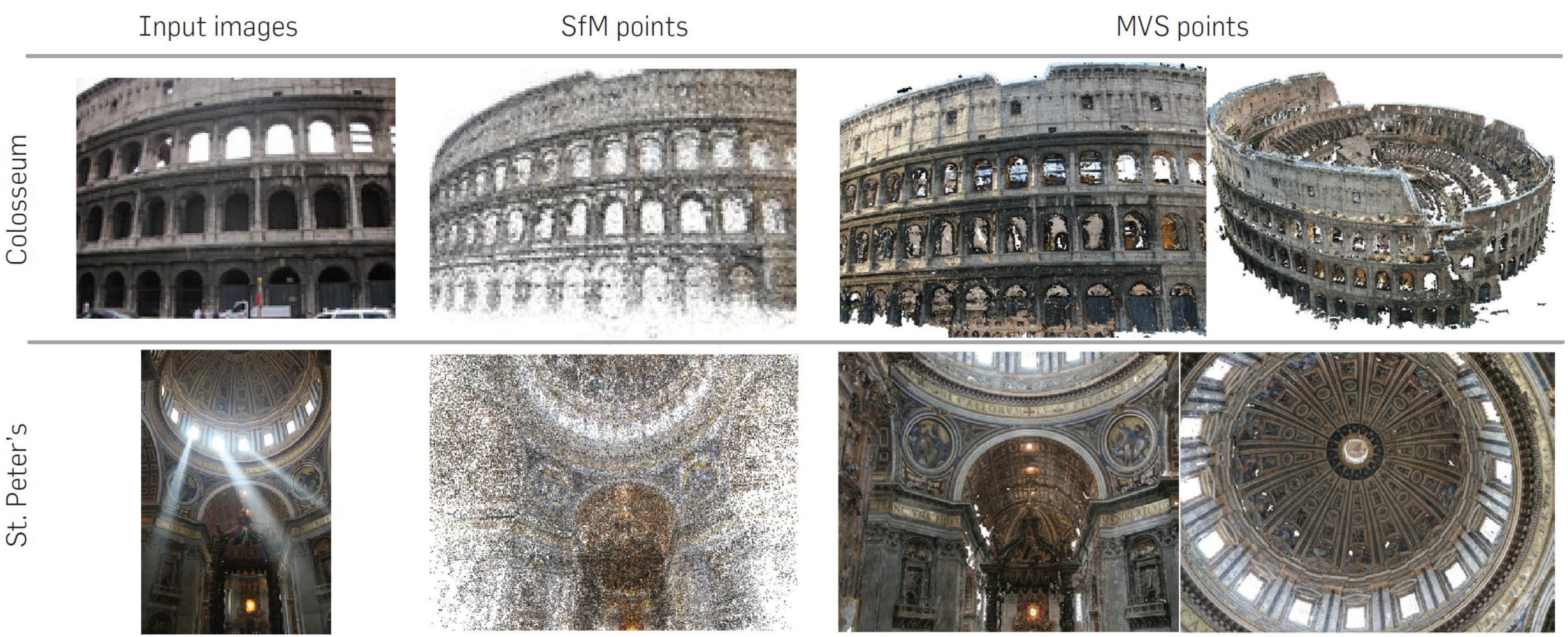
Variations on the Hermann grid: an extinction illusion

Jacques Ninio
Laboratoire de Physique Statistique[1], École Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France; e-mail: jacques.ninio@lps.ens.fr
Kent A Stevens
Department of Computer Science, Deschutes Hall, University of Oregon, Eugene, OR 97403, USA; e-mail: kent@cs.uoregon.edu
Received 21 September 1999, in revised form 21 June 2000

Abstract. When the white disks in a scintillating grid are reduced in size, and outlined in black, they tend to disappear. One sees only a few of them at a time, in clusters which move erratically on the page. Where they are not seen, the grey alleys seem to be continuous, generating grey crossings that are not actually present. Some black sparkling can be seen at those crossings where no disk is seen. The illusion also works in reverse contrast.

The Hermann grid (Brewster 1844; Hermann 1870) is a robust illusion. It is classically presented as a two-dimensional array of black squares, separated by rectilinear alleys. It is thought to be caused by processes of local brightness computation in arrays of

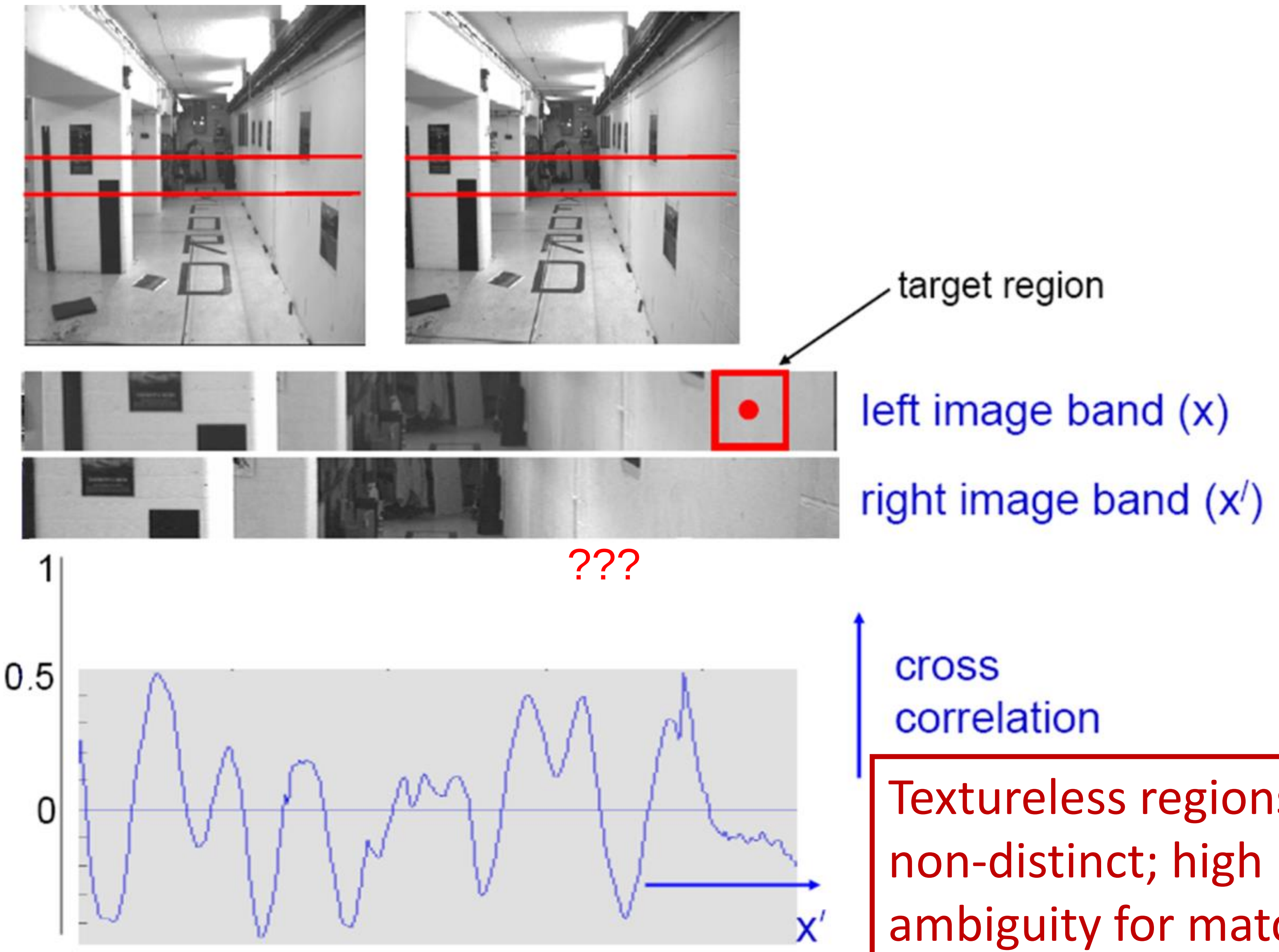
Sparse to Dense Correspodence



Building Rome in a Day

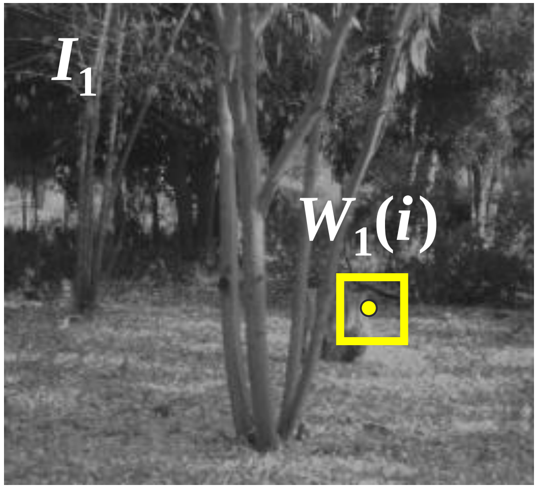
By Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, Richard Szeliski
Communications of the ACM, Vol. 54 No. 10, Pages 105-112

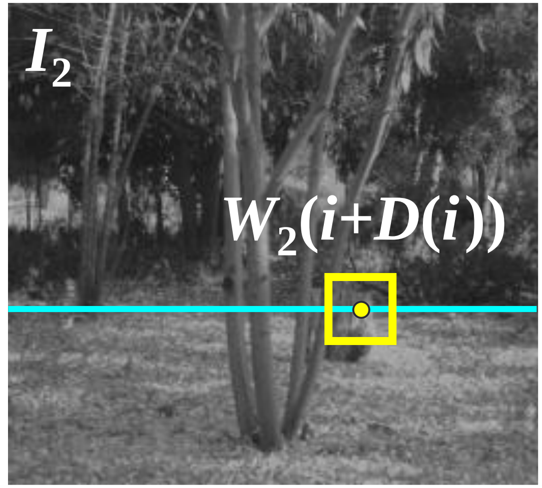
Correlation-based window matching

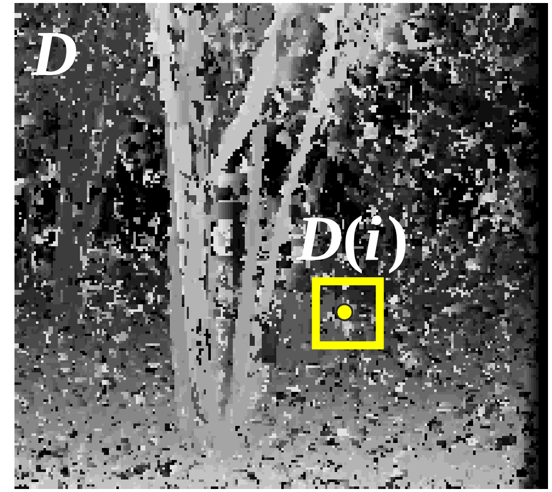


Textureless regions are non-distinct; high ambiguity for matches.

Stereo matching as energy minimization







$$E = \alpha E_{\text{data}}(I_1, I_2, D) + \beta E_{\text{smooth}}(D)$$

$$E_{\text{data}} = \sum_i \left(W_1(i) - W_2(i + D(i))\right)^2$$

$$E_{\text{smooth}} = \sum_{\text{neighbors}\, i,j} \rho\left(D(i) - D(j)\right)$$

- Energy functions of this form can be minimized using *graph cuts*

Y. Boykov, O. Veksler, and R. Zabih, [Fast Approximate Energy Minimization via Graph Cuts](), PAMI 2001

Source: Steve Seitz

Better results…

Graph cut method

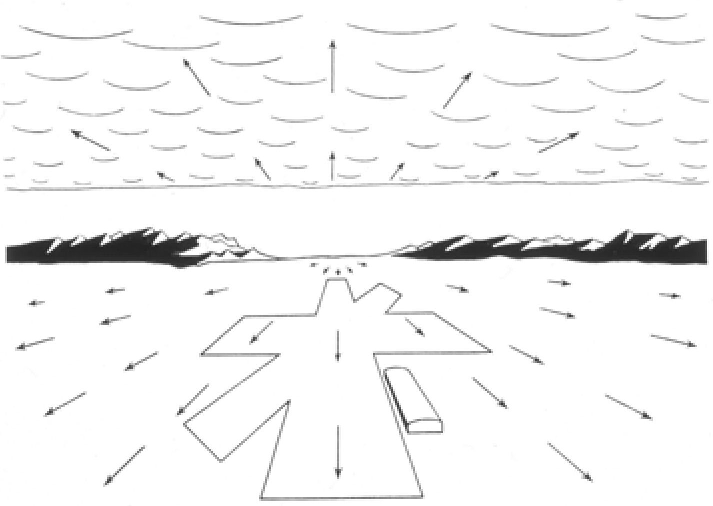
Ground truth

Boykov et al., [Fast Approximate Energy Minimization via Graph Cuts](), International Conference on Computer Vision, September 1999.

For the latest and greatest: http://www.middlebury.edu/stereo/

Computer Vision

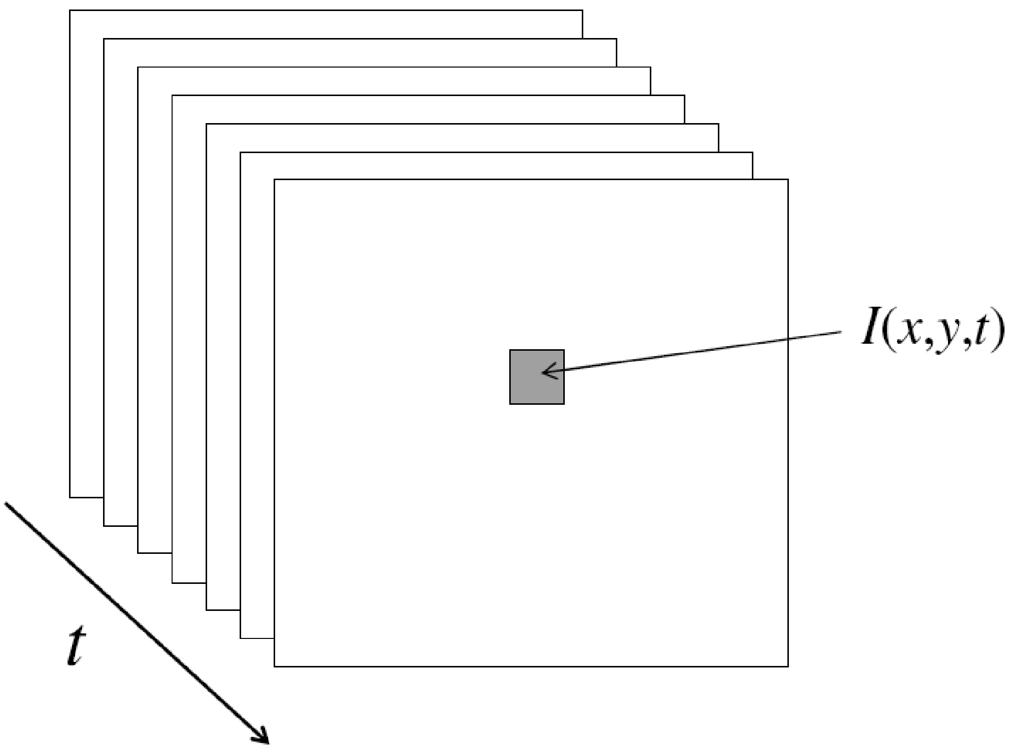
Motion and Optical Flow



Many slides adapted from S. Seitz, R. Szeliski, M. Pollefeys, K. Grauman and others…

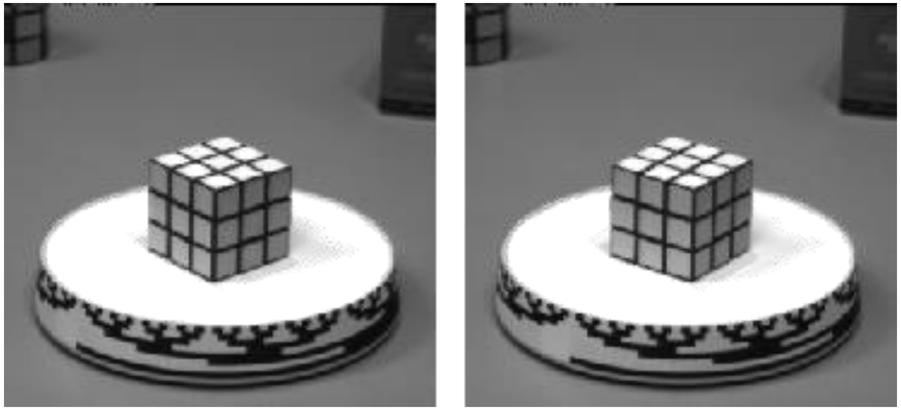
Video

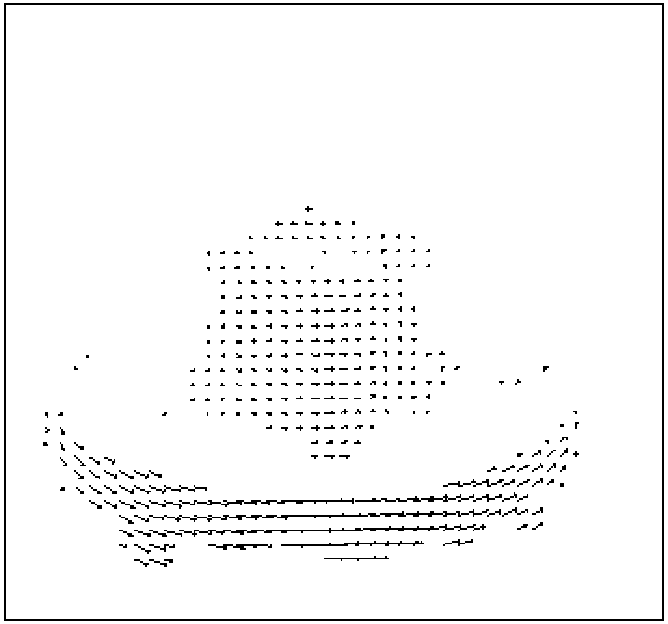
- A video is a sequence of frames captured over time
- Now our image data is a function of space (x, y) and time (t)



Motion estimation: Optical flow

Optic flow is the **apparent** motion of objects or surfaces

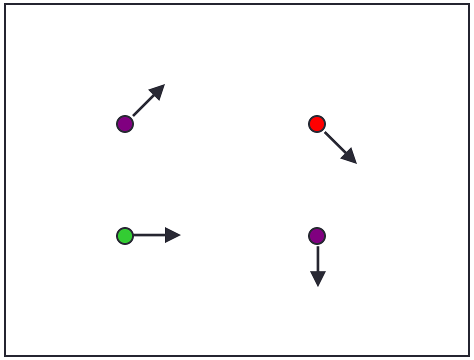




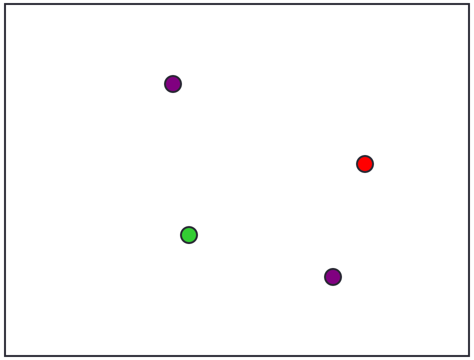
Will start by estimating motion of each pixel separately

Then will consider motion of entire image

Problem definition: optical flow



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



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

How to estimate pixel motion from image $I(x,y,t)$ to $I(x,y,t+1)$?

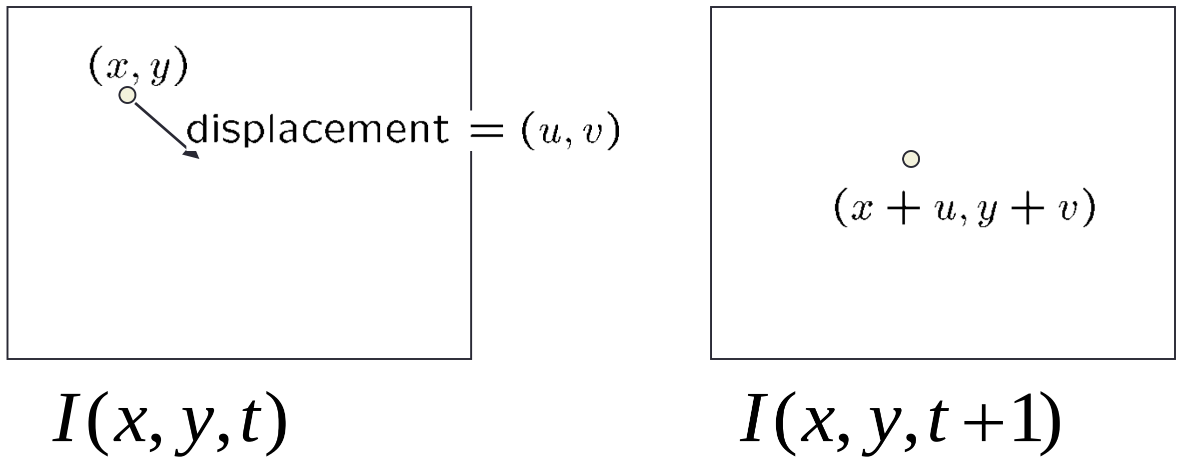
- Solve pixel correspondence problem
  - given a pixel in I(x,y,t), look for nearby pixels of the same color in *I(x,y,t+1)*

Key assumptions

- color constancy: a point in *I(x,y,* looks the same in I(x,y,t+1)
  - For grayscale images, this is brightness constancy
- small motion: points do not move very far

This is called the optical flow problem

Optical flow constraints (grayscale images)



- Let's look at these constraints more closely
  - brightness constancy constraint (equation)

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

  - small motion: (u and v are less than 1 pixel, or smooth)

    Taylor series expansion of I:

$$I(x+u, y+v) = I(x, y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + [\text{higher order terms}]$$

$$\approx I(x, y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v$$

Optical flow equation

- Combining these two equations

$$
\begin{aligned}
0 &= I(x+u, y+v, t+1) - I(x, y, t) \\
&\approx I(x, y, t+1) + I_x u + I_y v - I(x, y, t)
\end{aligned}
$$

(Short hand: $I_x = \frac{\partial I}{\partial x}$ for t **or** $t+1$)

Optical flow equation

- Combining these two equations

$$
\begin{aligned}
0 &= I(x+u, y+v, t+1) - I(x, y, t) \\
&\approx I(x, y, t+1) + I_x u + I_y v - I(x, y, t) \\
&\approx [I(x, y, t+1) - I(x, y, t)] + I_x u + I_y v \\
&\approx I_t + I_x u + I_y v \\
&\approx I_t + \nabla I \cdot < u, v >
\end{aligned}
$$

(Short hand: $I_x = \frac{\partial I}{\partial x}$ for t **or** $t+1$)

Optical flow equation

- Combining these two equations

$$
\begin{aligned}
0 &= I(x+u, y+v, t+1) - I(x, y, t) \\
&\approx I(x, y, t+1) + I_x u + I_y v - I(x, y, t) \\
&\approx \boxed{[I(x, y, t+1) - I(x, y, t)]} + I_x u + I_y v \\
&\approx \boxed{I_t} + I_x u + I_y v \\
&\approx I_t + \nabla I \cdot \langle u, v \rangle
\end{aligned}
$$

(Short hand: $I_x = \frac{\partial I}{\partial x}$ for t **or** $t+1$)

In the limit as u and v go to zero, this becomes exact

$$0 = I_t + \nabla I \cdot \langle u, v \rangle$$

Brightness constancy constraint equation

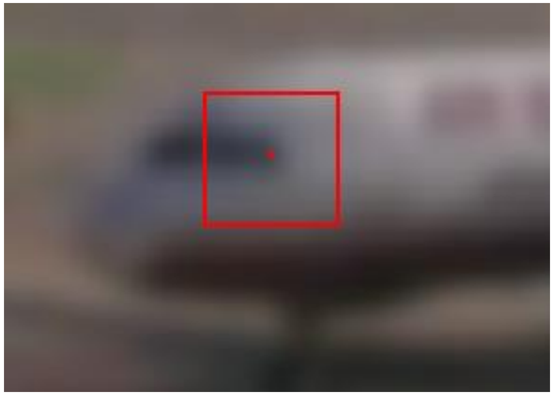
$$I_x u + I_y v + I_t = 0$$

How does this make sense?

Brightness constancy constraint equation

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

- What do the static image gradients have to do with motion estimation?



The brightness constancy constraint

Can we use this equation to recover image motion (u,v) at each pixel?

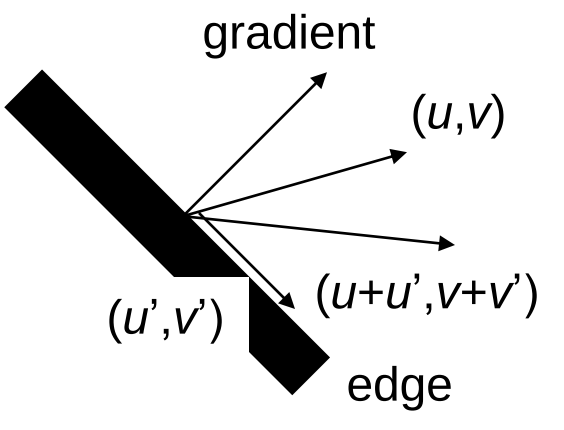
$$0 = I_t + \nabla I \cdot < u, v > \quad \text{or} \quad I_x u + I_y v + I_t = 0$$

- How many equations and unknowns per pixel?
  - One equation (this is a scalar equation!), two unknowns (u,v)

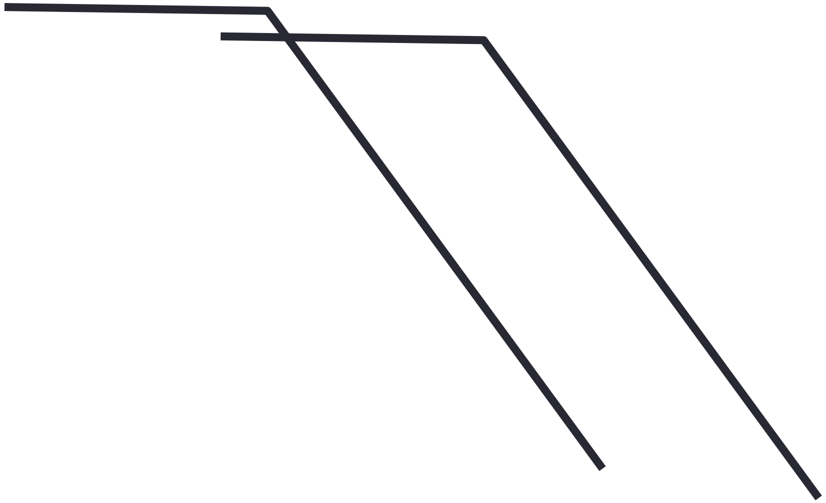
The component of the motion perpendicular to the gradient (i.e., parallel to the edge) cannot be measured

If (*u*, *v*) satisfies the equation, so does (*u*+*u'*, *v*+*v'*) if

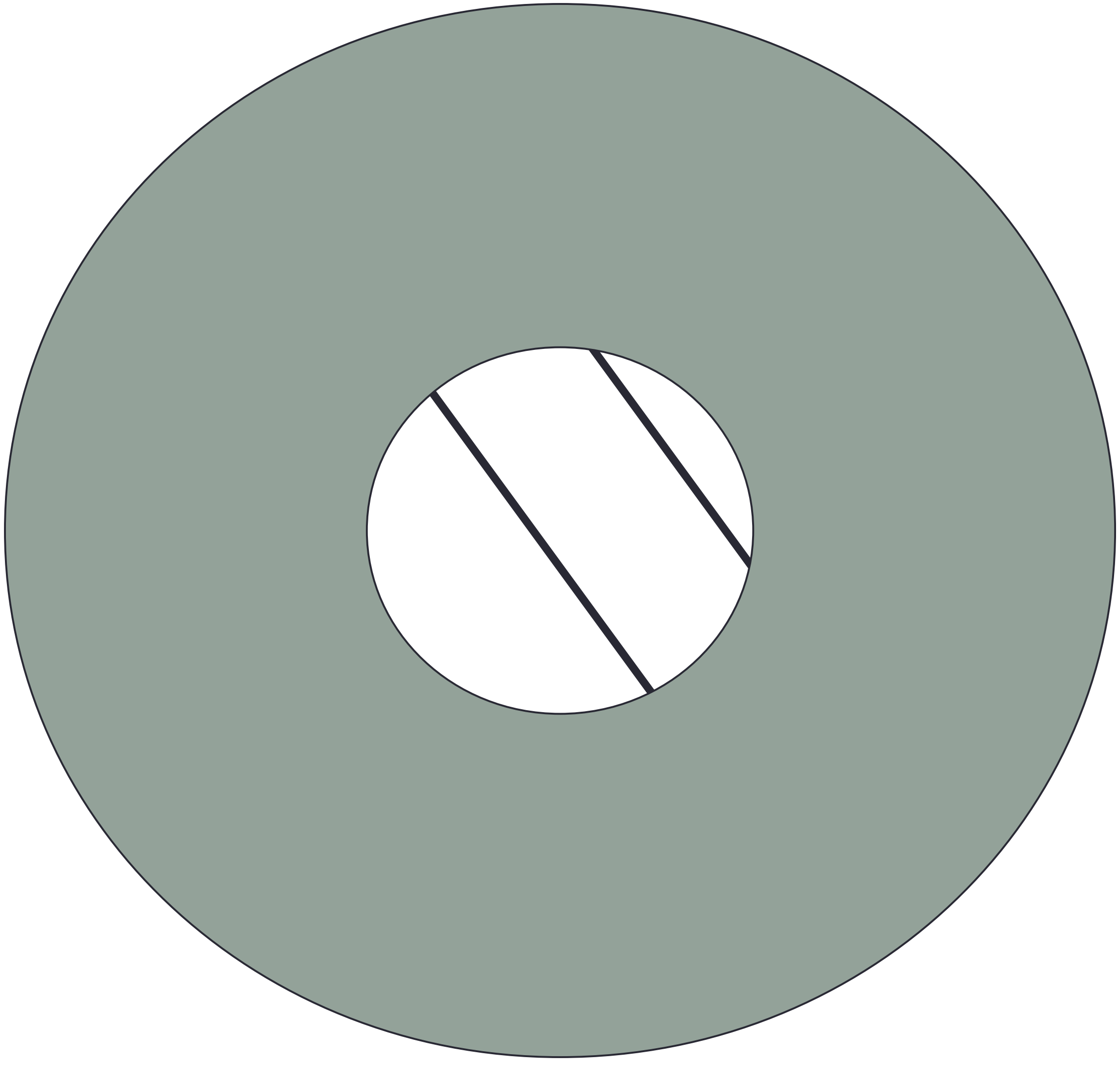
$$\nabla \mathrm{I} \cdot [\mathrm{u'} \;\; \mathrm{v'}]^{\mathrm{T}} = 0$$



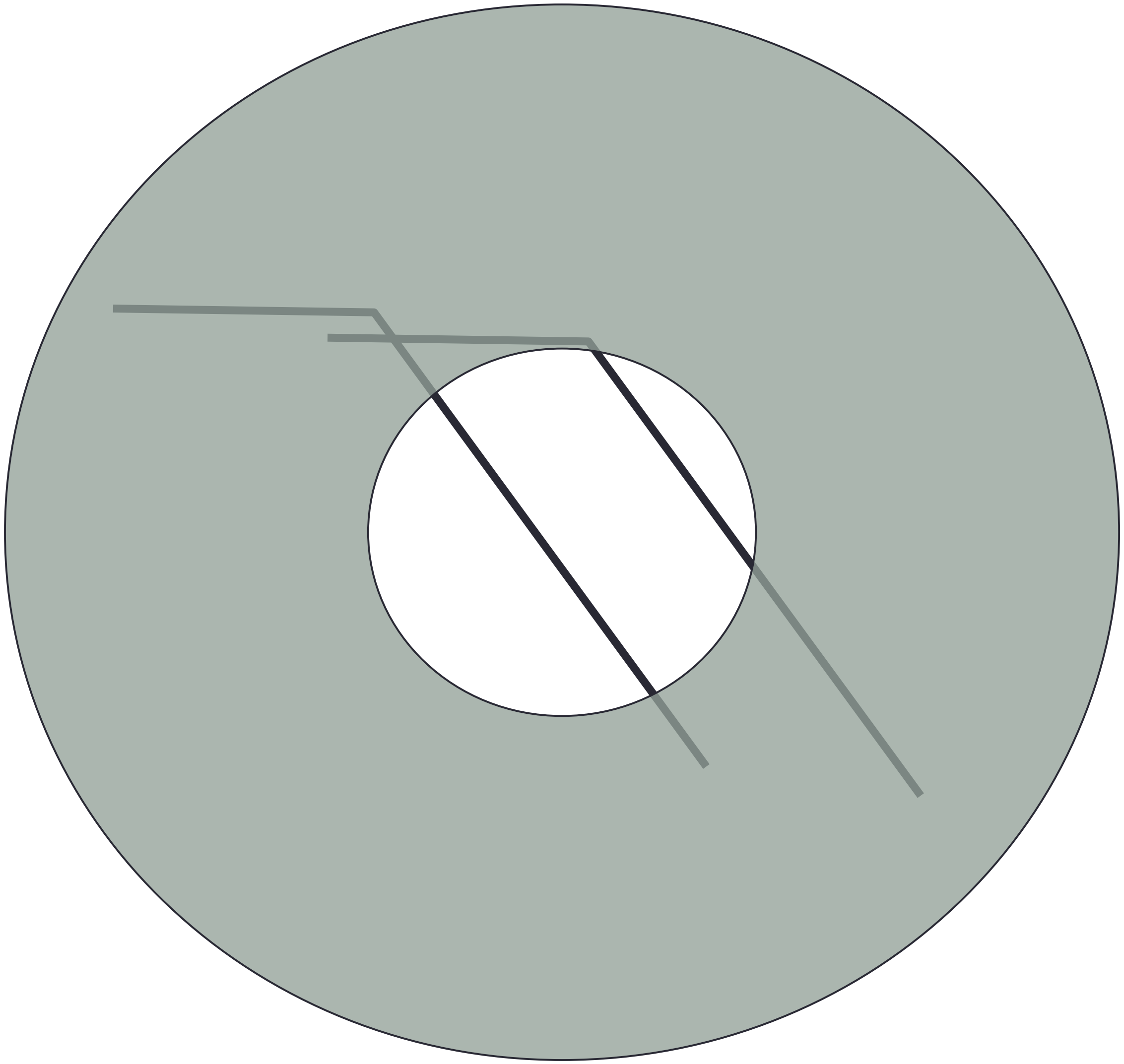
Aperture problem



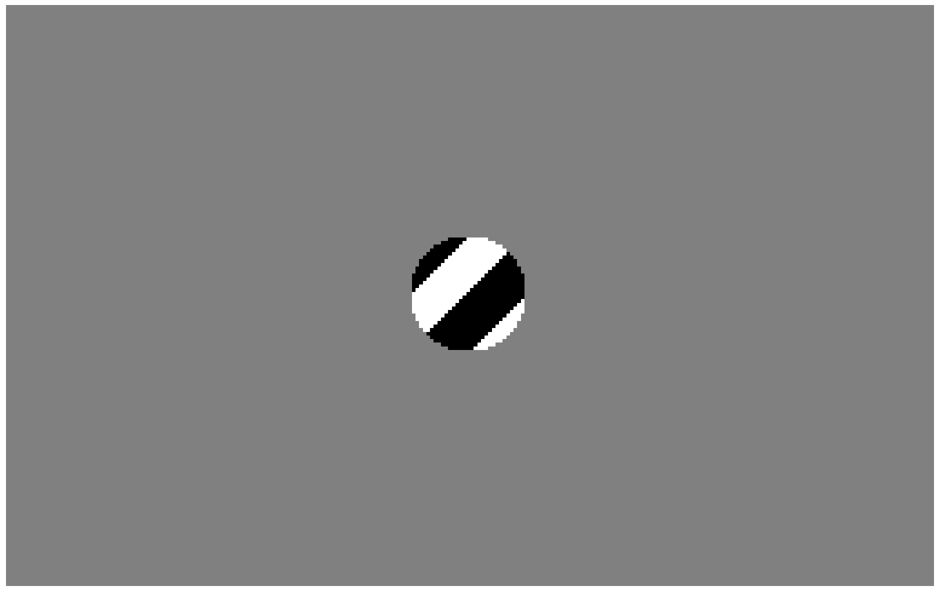
Aperture problem



Aperture problem

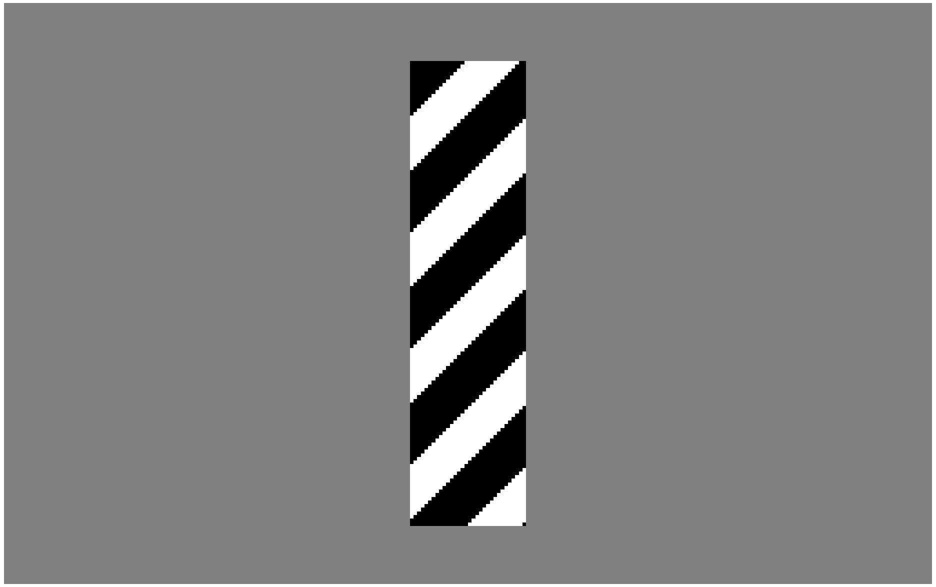


The barber pole illusion



http://en.wikipedia.org/wiki/Barberpole_illusion

The barber pole illusion



http://en.wikipedia.org/wiki/Barberpole_illusion

Solving the ambiguity…

B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 674–679, 1981.

- How to get more equations for a pixel?
- **Spatial coherence constraint**
- Assume the pixel's neighbors have the same (u,v)
  - If we use a 5x5 window, that gives us 25 equations per pixel

$$0 = I_t(\mathbf{p_i}) + \nabla I(\mathbf{p_i}) \cdot [u \ \ v]$$

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

Solving the ambiguity…

- Least squares problem:

$$\begin{bmatrix} I_x(\mathbf{p_1}) & I_y(\mathbf{p_1}) \\ I_x(\mathbf{p_2}) & I_y(\mathbf{p_2}) \\ \vdots & \vdots \\ I_x(\mathbf{p_{25}}) & I_y(\mathbf{p_{25}}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p_1}) \\ I_t(\mathbf{p_2}) \\ \vdots \\ I_t(\mathbf{p_{25}}) \end{bmatrix} \qquad \underset{25\times2}{A} \; \underset{2\times1}{d} = \underset{25\times1}{b}$$

Matching patches across images

- Overconstrained linear system

$$\begin{bmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{25}) & I_y(\mathbf{p}_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{25}) \end{bmatrix} \qquad \underset{25\text{x}2}{A} \ \underset{2\text{x}1}{d} = \underset{25\text{x}1}{b}$$

Least squares solution for *d* given by $(A^T A)\ d = A^T b$

$$\underbrace{\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} u \\ v \end{bmatrix} = - \underbrace{\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}}_{A^T b}$$

The summations are over all pixels in the K x K window

Conditions for solvability

Optimal (u, v) satisfies Lucas-Kanade equation

$$\underbrace{\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} u \\ v \end{bmatrix} = - \underbrace{\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}}_{A^T b}$$

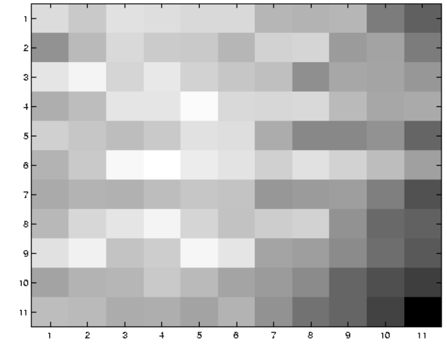
When is this solvable? I.e., what are good points to track?

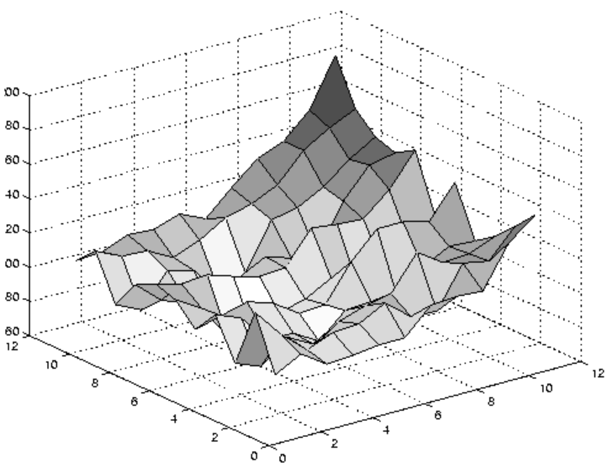
- **A^TA** should be invertible
- **A^TA** should not be too small due to noise
  - eigenvalues λ_1 and λ_2 of **A^TA** should not be too small
- **A^TA** should be well-conditioned
  - λ_1 / λ_2 should not be too large (λ_1 = larger eigenvalue)

Does this remind you of anything?

Criteria for Harris corner detector

Low texture region



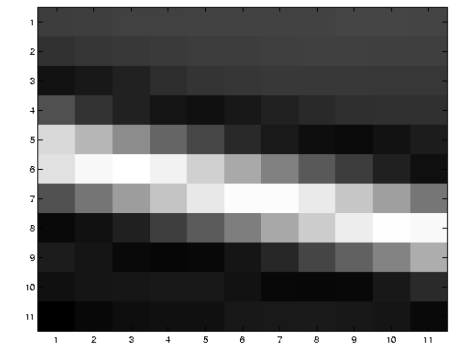


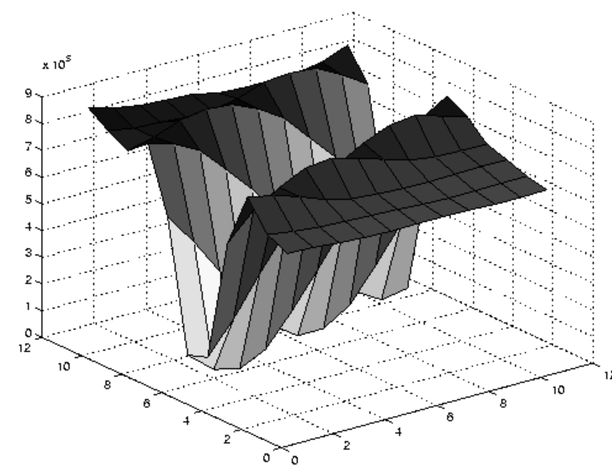
$$\sum \nabla I (\nabla I)^T$$

– gradients have small magnitude

– small λ_1, small λ_2

Edge

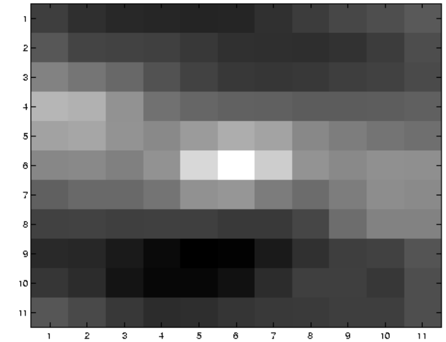


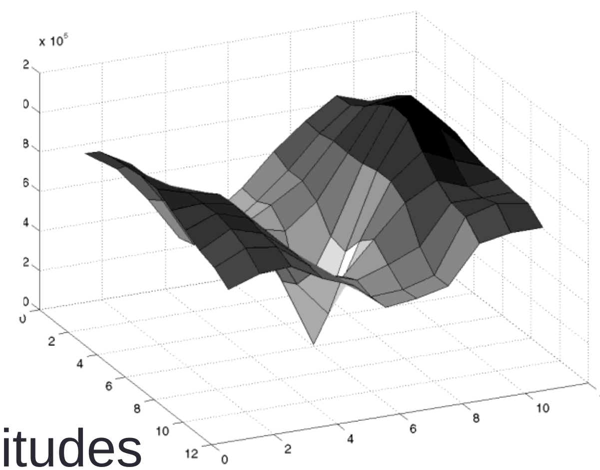


$$\sum \nabla I (\nabla I)^T$$

- large gradients, all the same
- large λ_1, small λ_2

High textured region



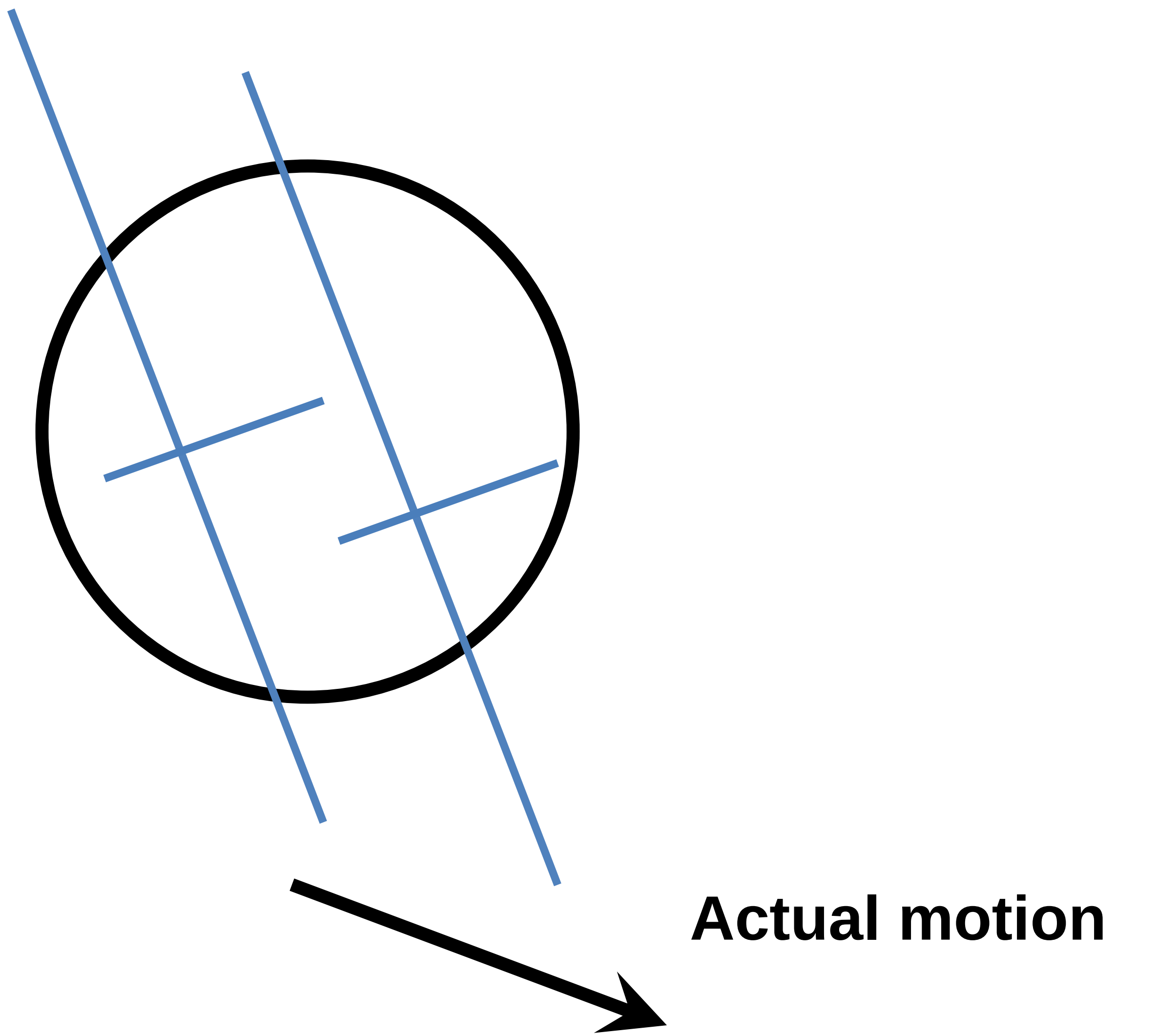


$$\sum \nabla I (\nabla I)^T$$

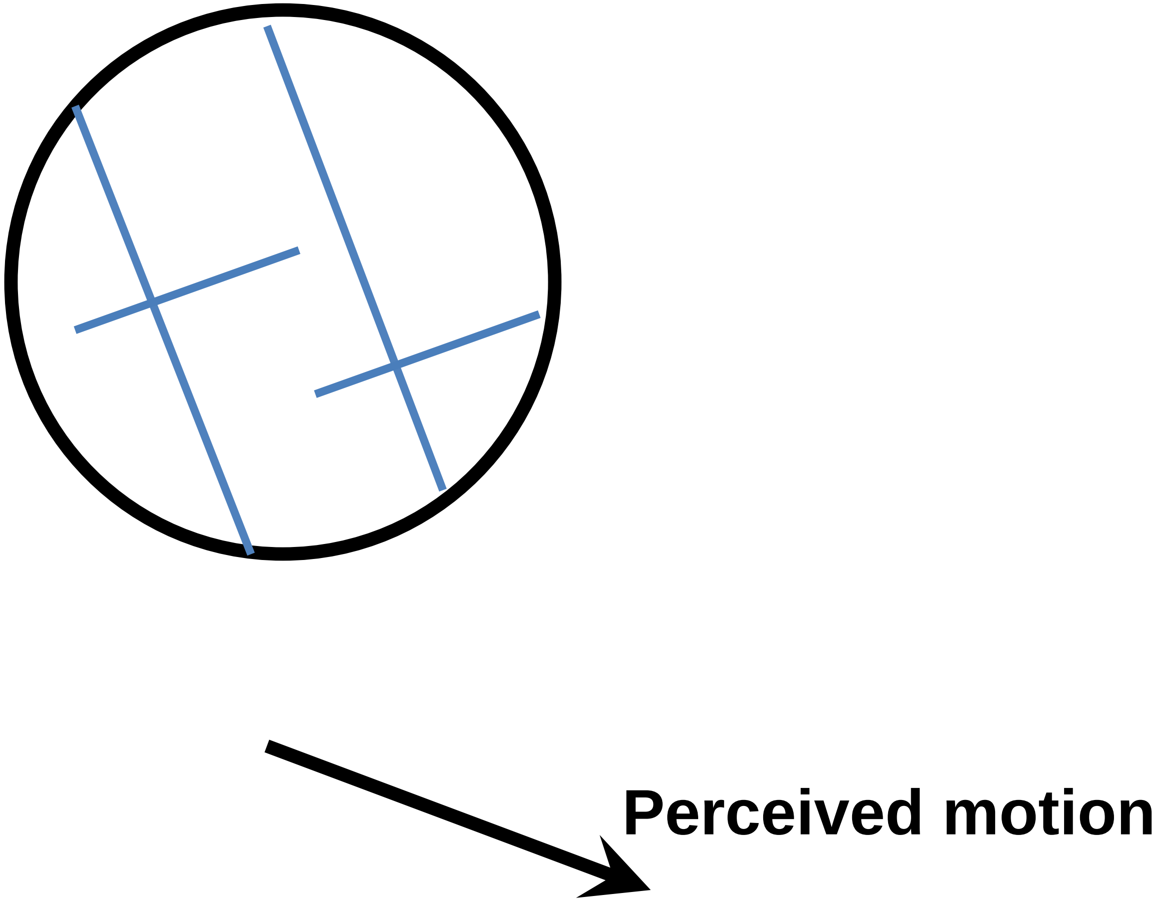
– gradients are different, large magnitudes

– large λ_1, large λ_2

The aperture problem resolved



The aperture problem resolved



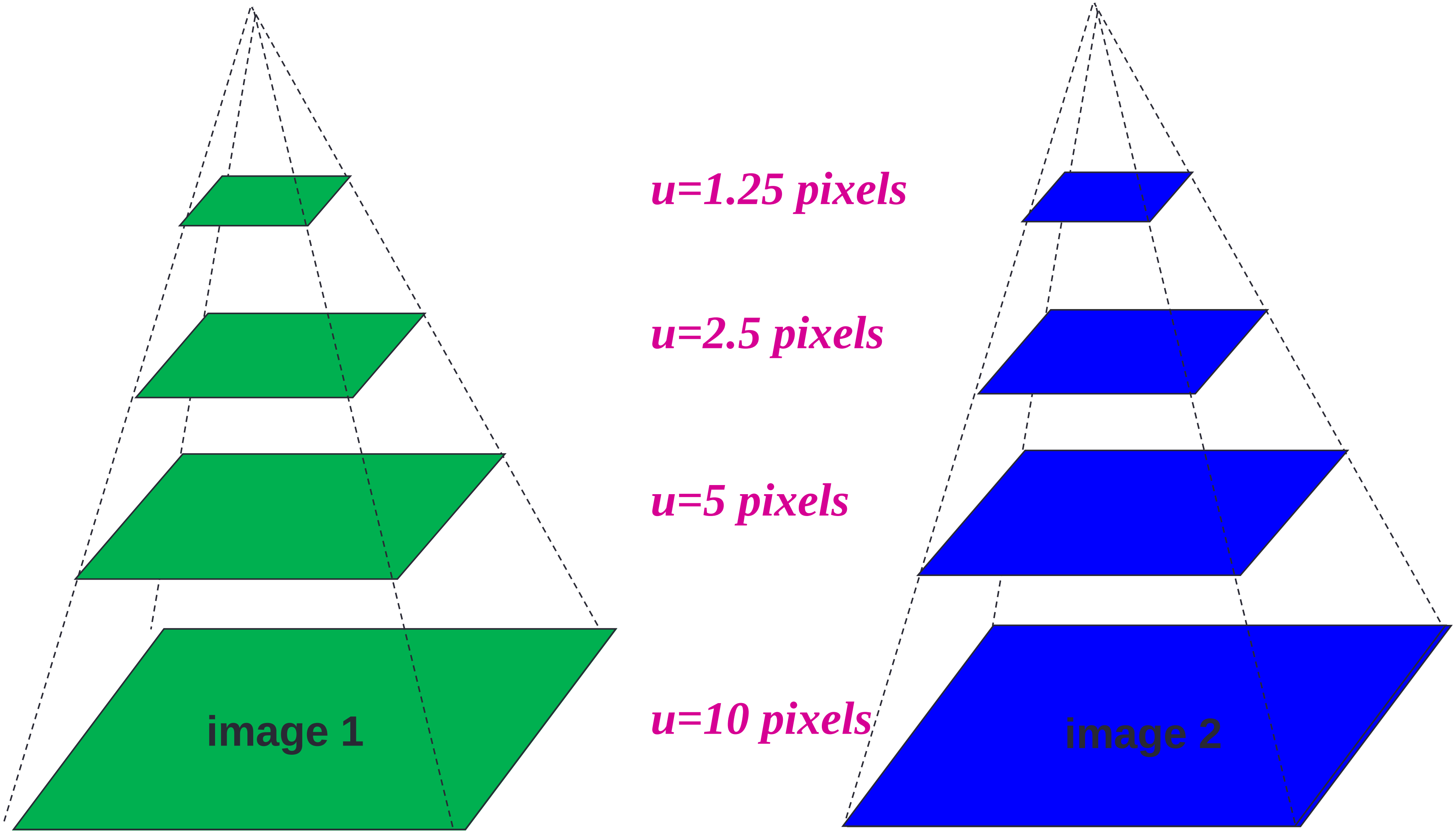
Errors in Lucas-Kanade

- A point does not move like its neighbors
  - Motion segmentation
- Brightness constancy does not hold
  - Do exhaustive neighborhood search with normalized correlation - tracking features – maybe SIFT – more later….
- The motion is large (larger than a pixel)
  1. Not-linear: Iterative refinement
  2. Local minima: coarse-to-fine estimation

Revisiting the small motion assumption

- Is this motion small enough?
  - Probably not—it's much larger than one pixel
  - How might we solve this problem?

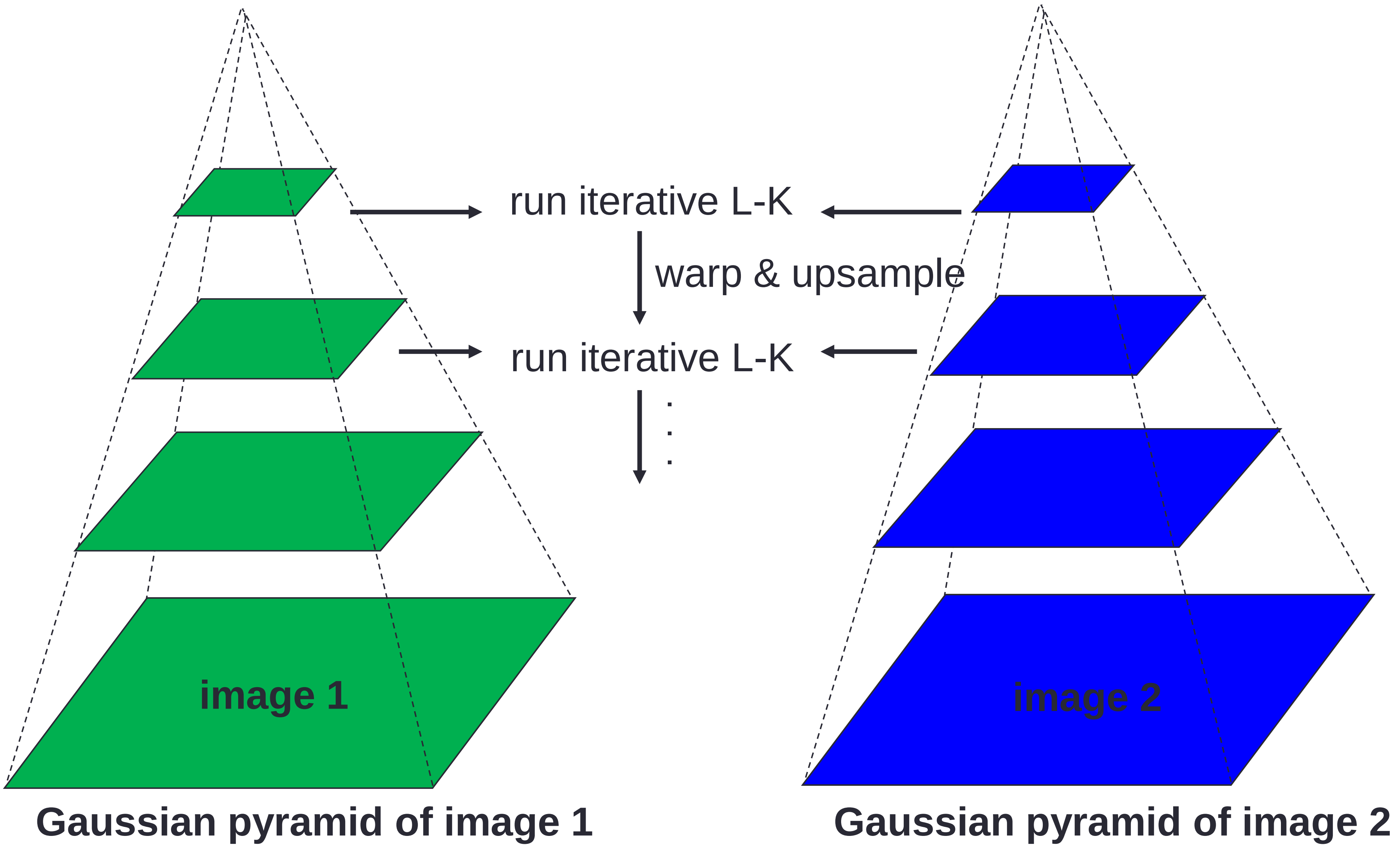
Coarse-to-fine optical flow estimation



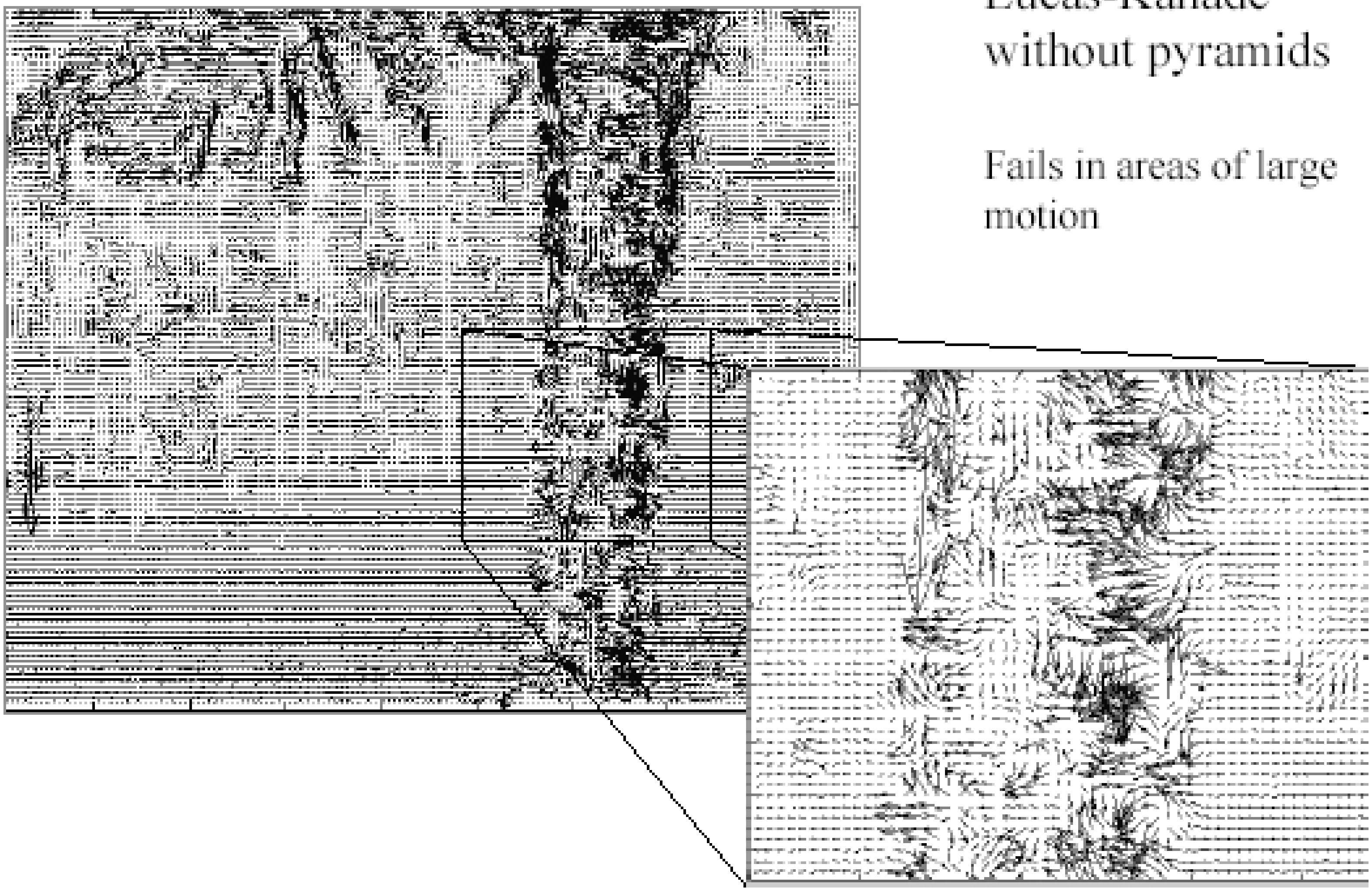
Gaussian pyramid of image 1

Gaussian pyramid of image 2

Coarse-to-fine optical flow estimation



Optical Flow Results

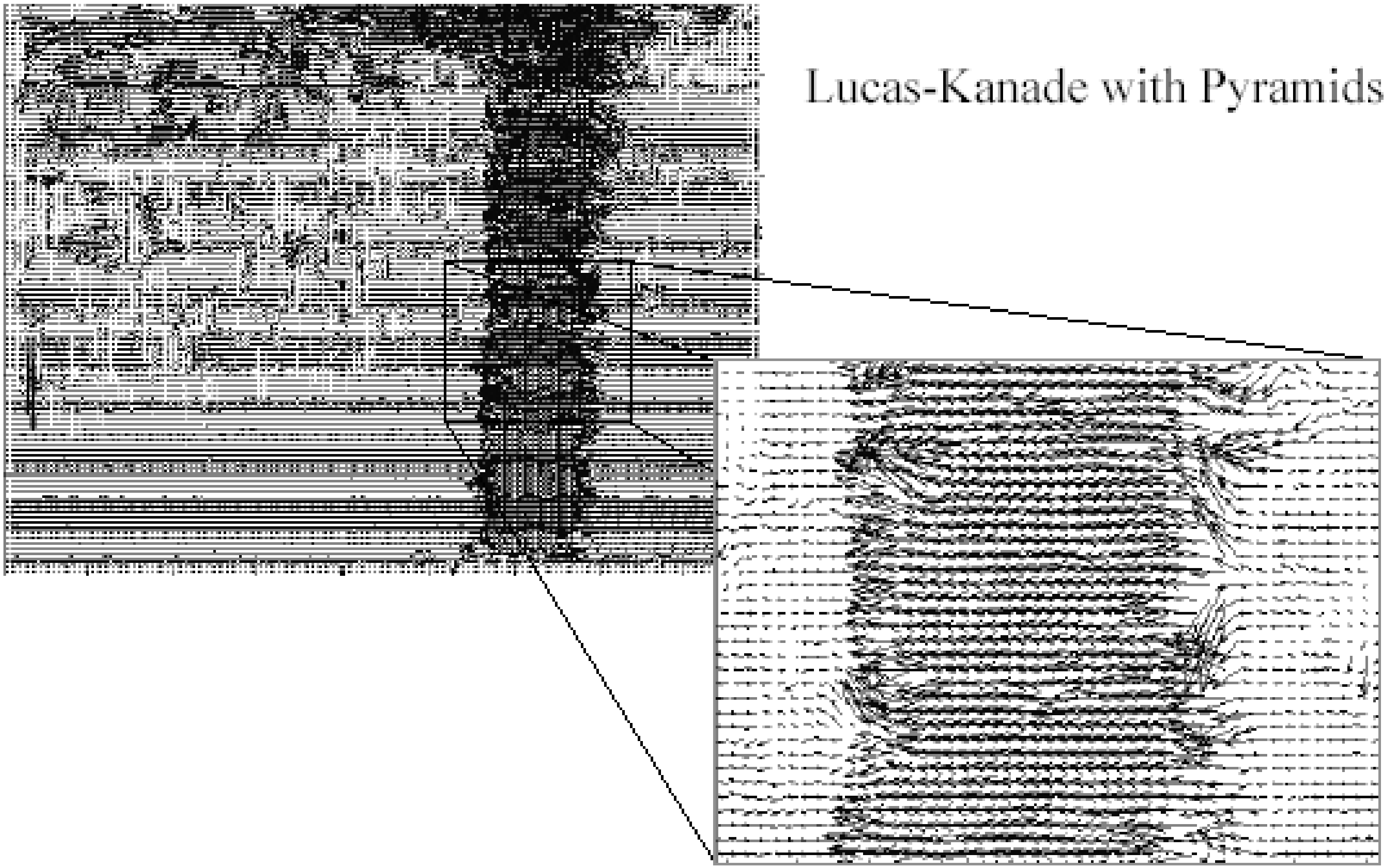


Lucas-Kanade
without pyramids

Fails in areas of large
motion

* From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

Optical Flow Results



Lucas-Kanade with Pyramids

* From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

State-of-the-art optical flow, 2009

Start with something similar to Lucas-Kanade

+ gradient constancy

+ energy minimization with smoothing term

+ region matching

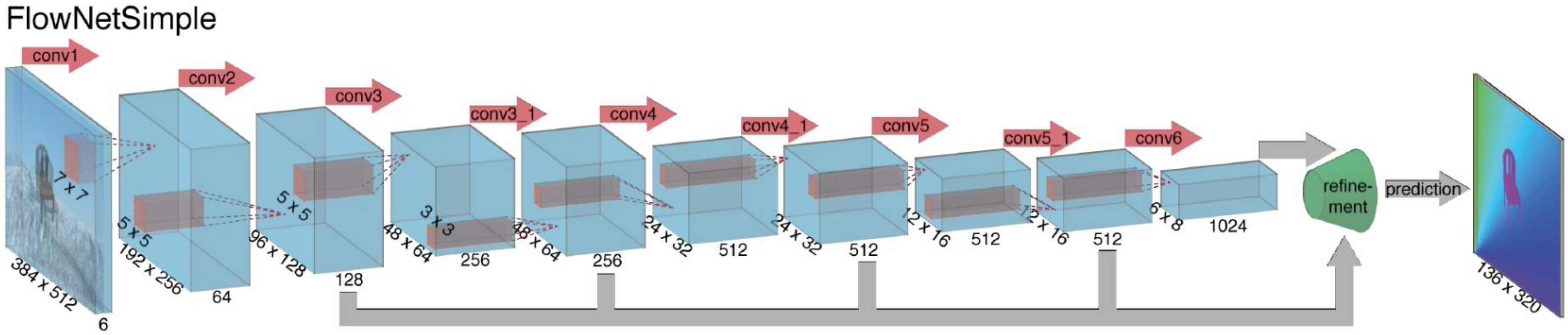
+ keypoint matching (long-range)

Region-based +Pixel-based +Keypoint-based

Large displacement optical flow, Brox et al., CVPR 2009

State-of-the-art optical flow, 2015

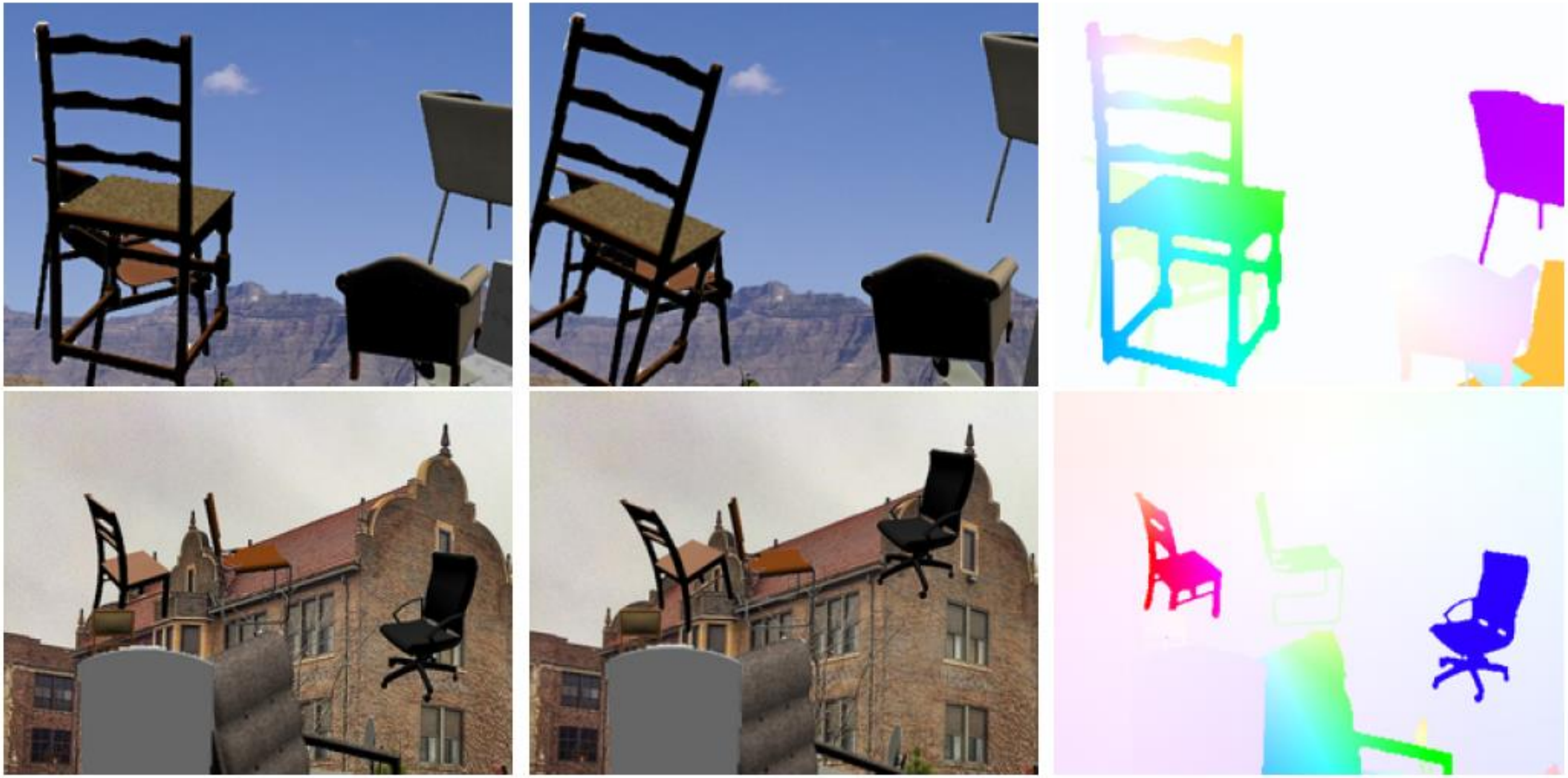
Deep convolutional network which accepts a pair of input frames and upsamples the estimated flow back to input resolution. Very fast because of deep network, near the state-of-the-art in terms of end-point-error.



Fischer et al. 2015. https://arxiv.org/abs/1504.06852

Deep optical flow, 2015

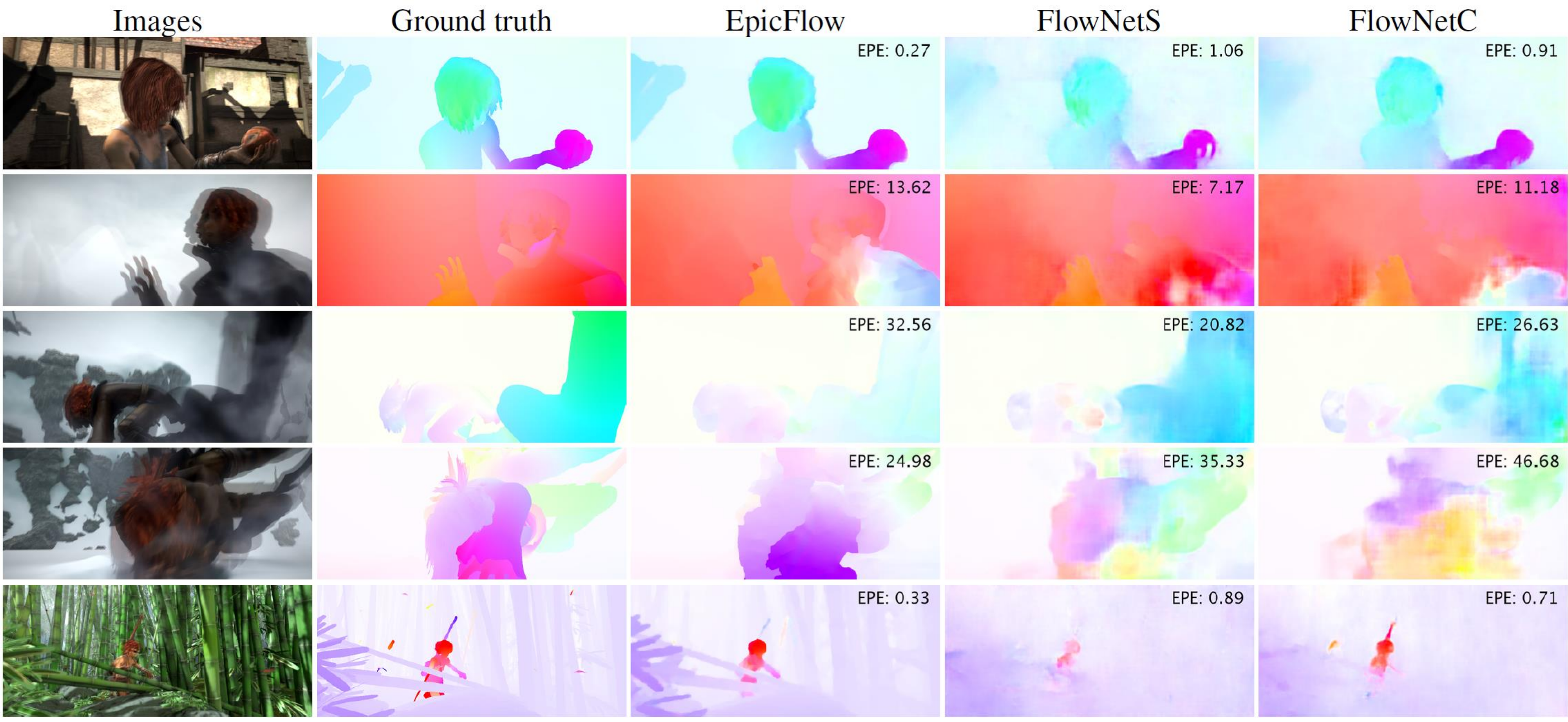
Synthetic Training data



Fischer et al. 2015. https://arxiv.org/abs/1504.06852

Deep optical flow, 2015

Results on Sintel



Fischer et al. 2015. https://arxiv.org/abs/1504.06852

Optical flow

- Definition: optical flow is the *apparent* motion of brightness patterns in the image
- Ideally, optical flow would be the same as the motion field
- Have to be careful: apparent motion can be caused by lighting changes without any actual motion
  - Think of a uniform rotating sphere under fixed lighting vs. a stationary sphere under moving illumination