Motion and Optical Flow

Ali Farhadi

CSE 576

We live in a moving world

- Perceiving, understanding and predicting motion is an important part of our daily lives

Motion and perceptual organization

- Even "impoverished" motion data can evoke a strong percept

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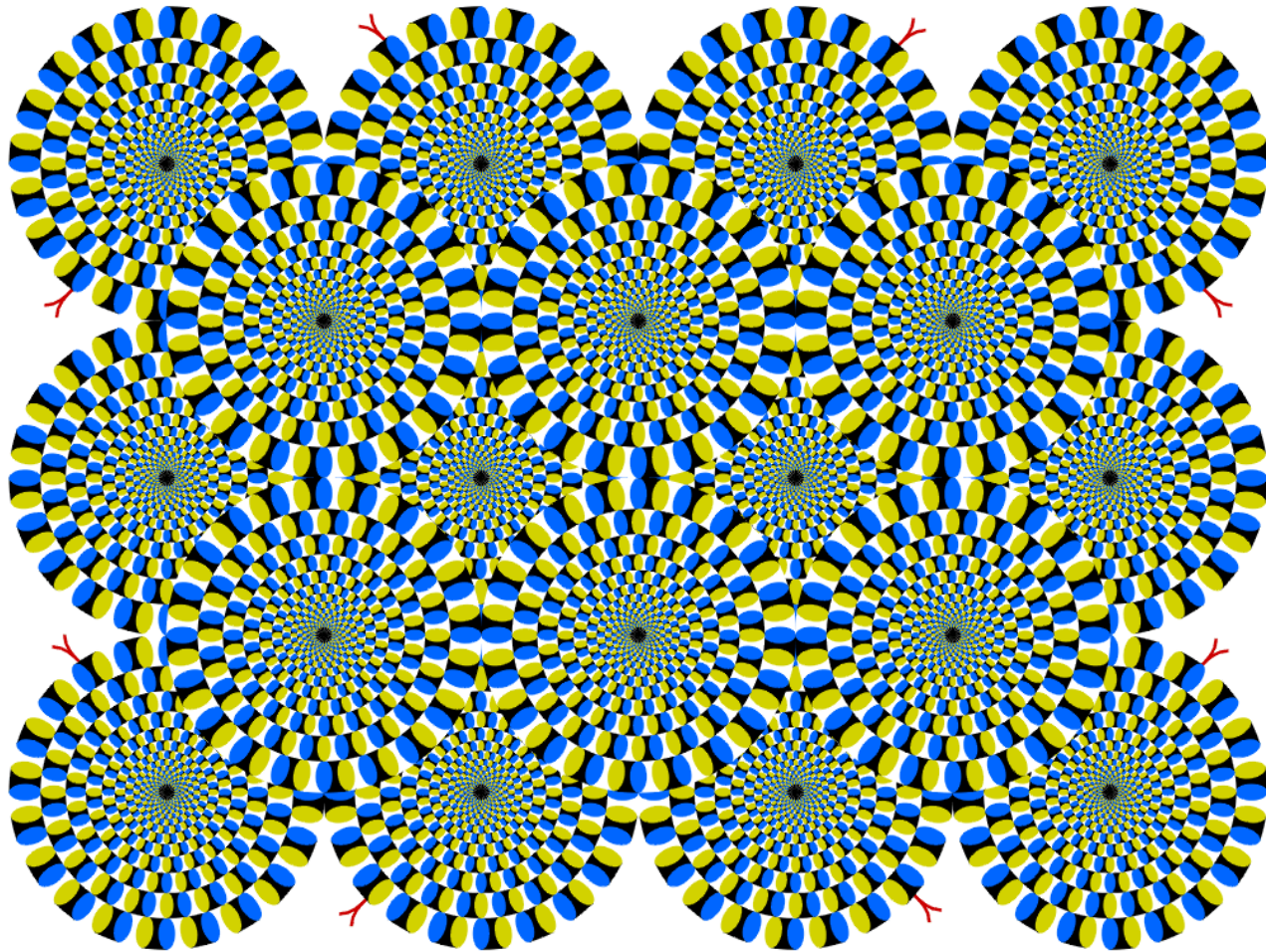
Motion and perceptual organization

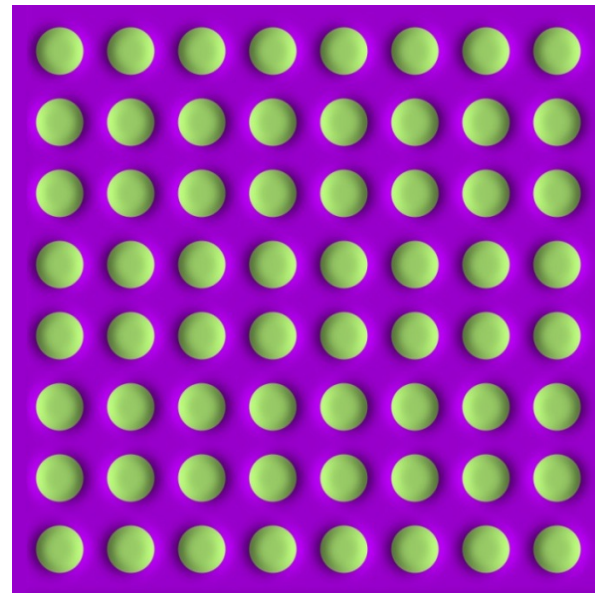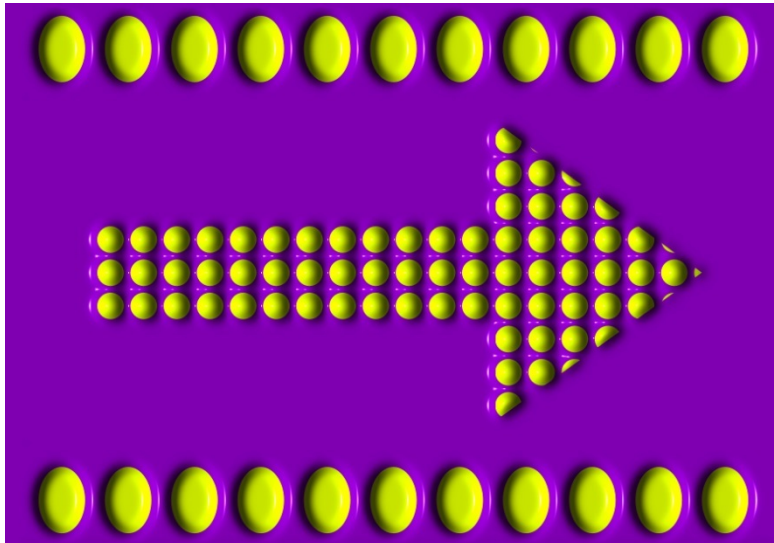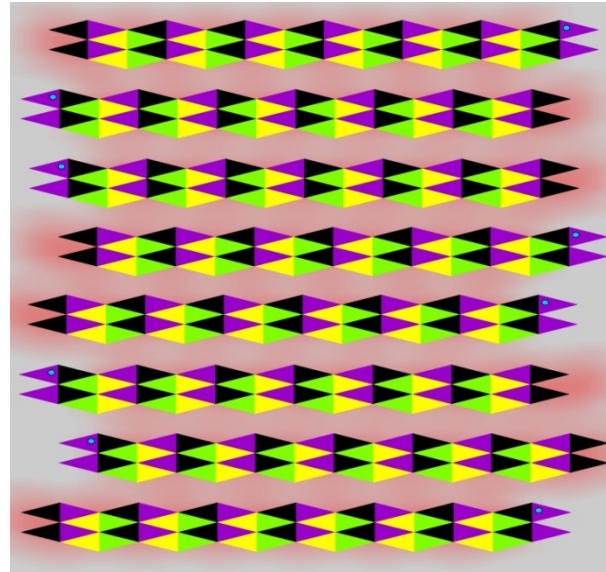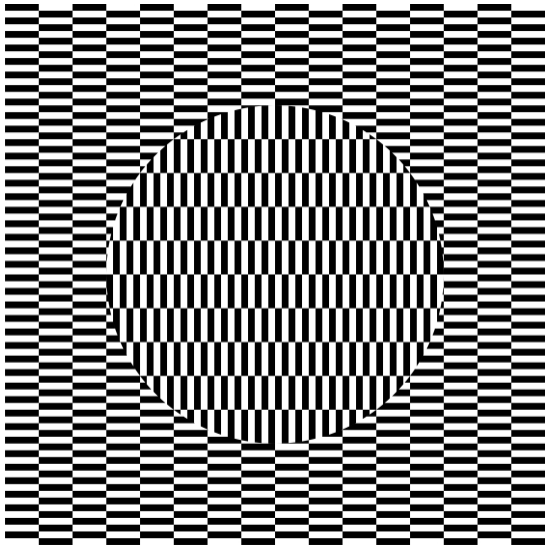- Even "impoverished" motion data can evoke a strong percept

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

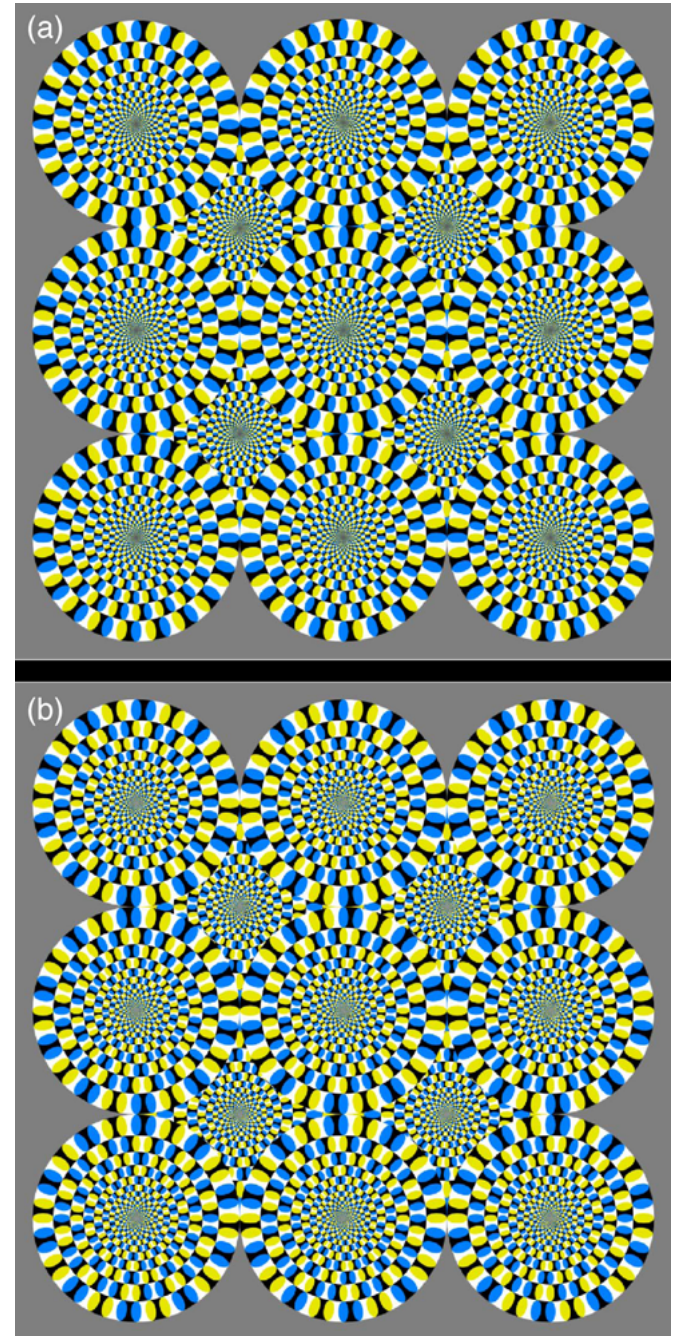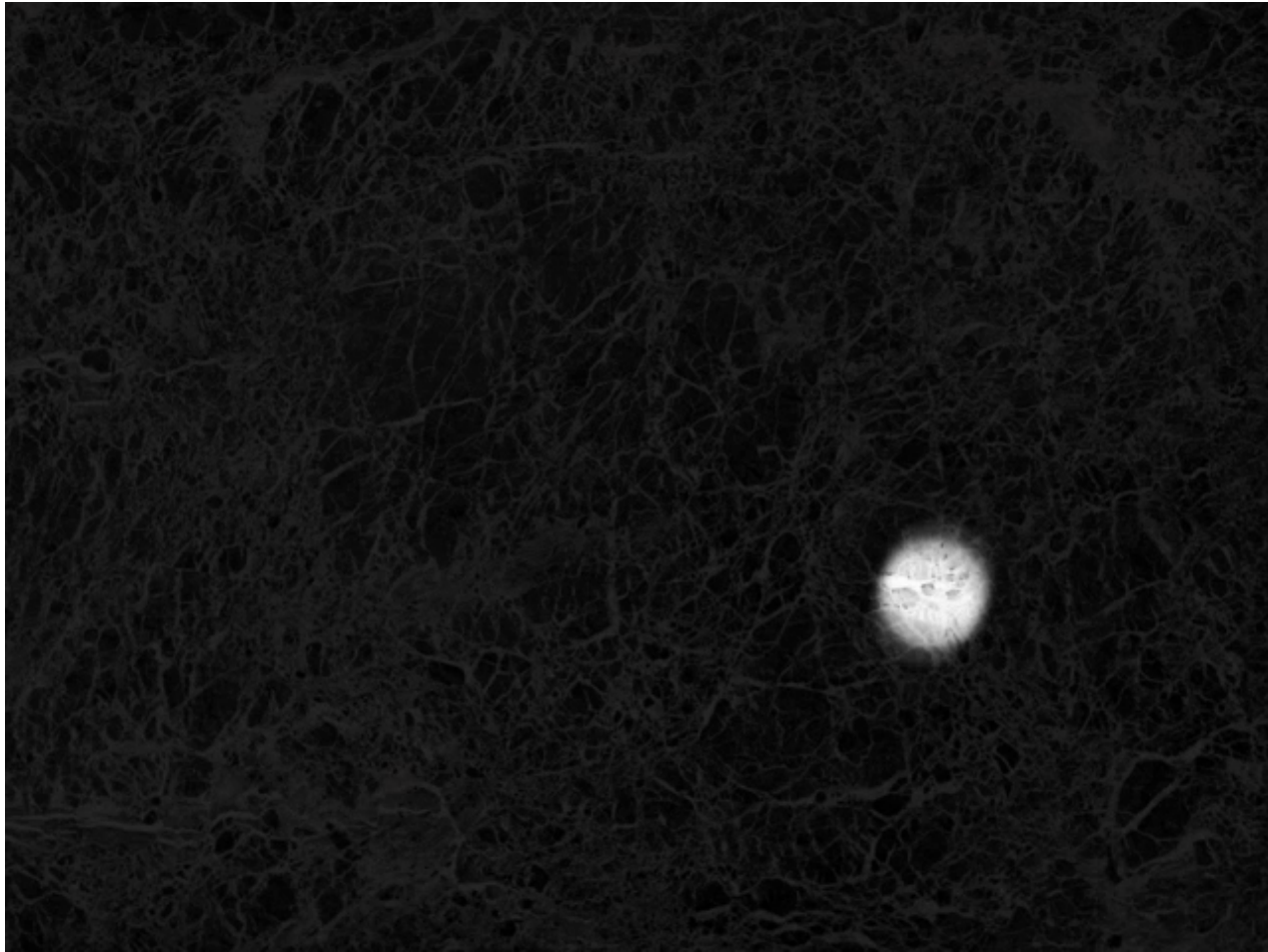# Seeing motion from a static picture?
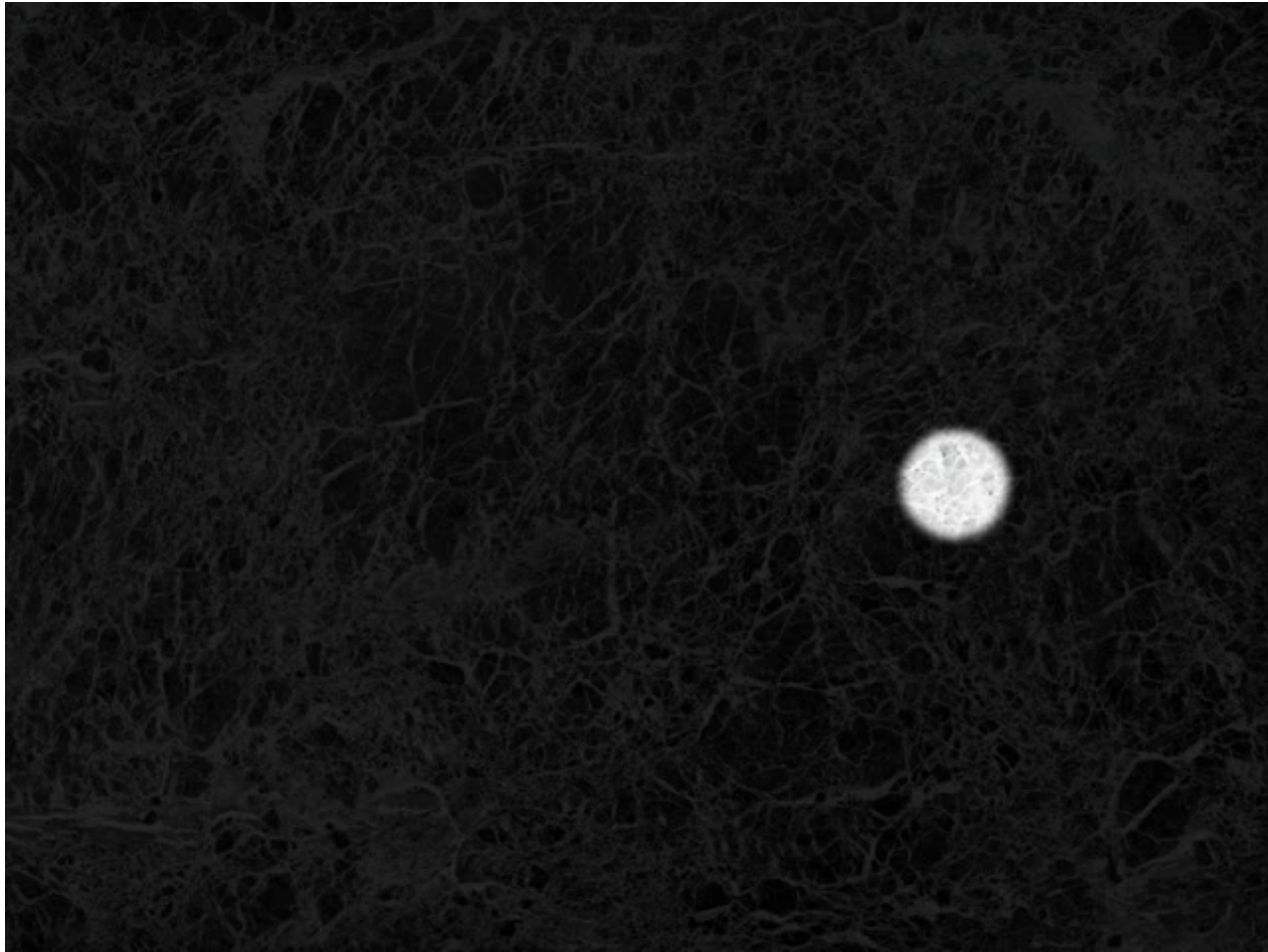
More examples

How is this possible?

- The true mechanism is to be revealed

- FMRI data suggest that illusion is related to some component of eye movements

- We don't expect computer vision to "see" motion from these stimuli, yet

What do you see?

In fact, …

The cause of motion

- Three factors in imaging process
  - Light
  - Object
  - Camera
- Varying either of them causes motion
  - Static camera, moving objects (surveillance)
  - Moving camera, static scene (3D capture)
  - Moving camera, moving scene (sports, movie)
  - Static camera, moving objects, moving light (time lapse)
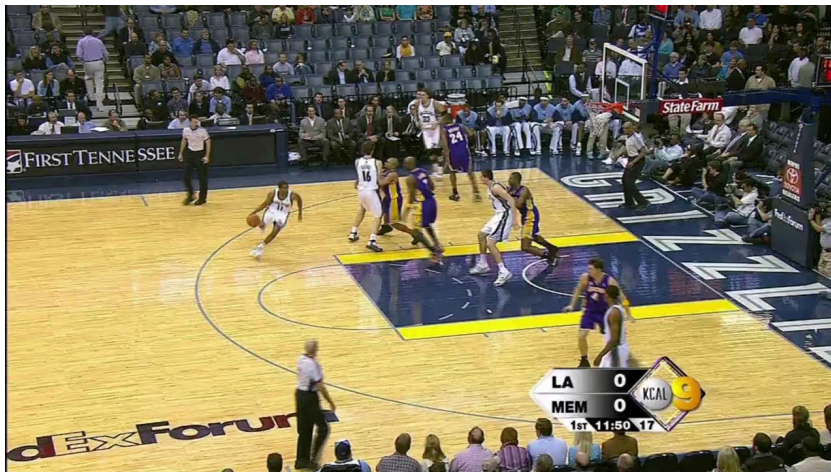
Motion scenarios (priors)

Static camera, moving scene

Moving camera, static scene

Moving camera, moving scene

Static camera, moving scene, moving light

We still don't touch these areas

How can we recover motion?

Recovering motion

- ## Feature-tracking
  - Extract visual features (corners, textured areas) and "track" them over multiple frames

- ## Optical flow
  - Recover image motion at each pixel from spatio-temporal image brightness variations (optical flow)

Two problems, one registration method

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.
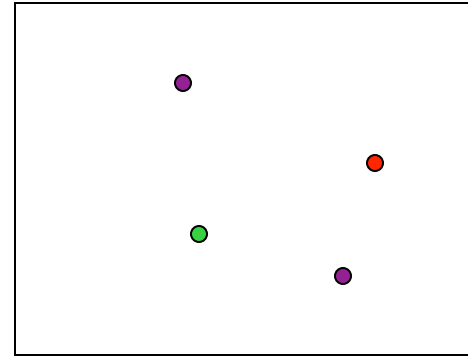
Feature tracking

- Challenges
  - Figure out which features can be tracked
  - Efficiently track across frames
  - Some points may change appearance over time (e.g., due to rotation, moving into shadows, etc.)
  - Drift: small errors can accumulate as appearance model is updated
  - Points may appear or disappear: need to be able to add/delete tracked points
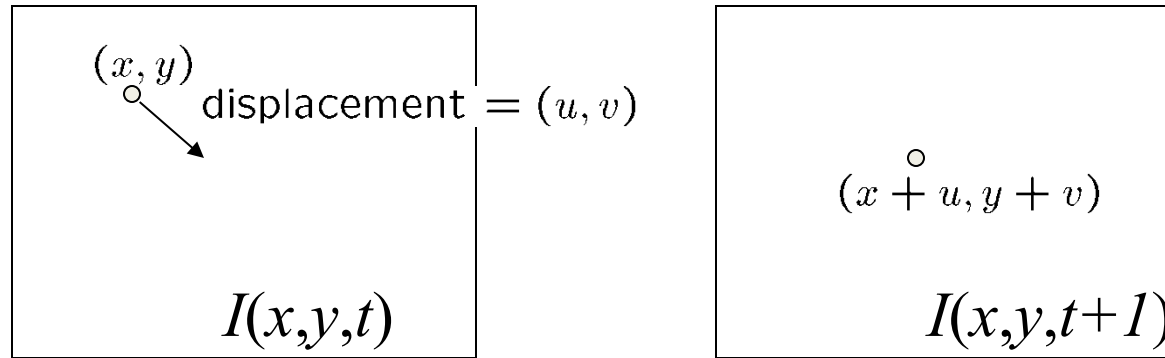
Feature tracking
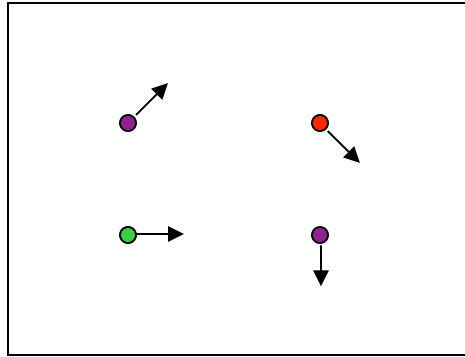
$I(x,y,t)$ $I(x,y,t+1)$

- Given two subsequent frames, estimate the point translation

- Key assumptions of Lucas-Kanade Tracker
  - **Brightness constancy:** projection of the same point looks the same in every frame
  - **Small motion:** points do not move very far
  - **Spatial coherence:** points move like their neighbors

The brightness constancy constraint

- Brightness Constancy Equation:

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

Take Taylor expansion of $I(x+u, y+v, t+1)$ at (x,y,t) to linearize the right side:

Image derivative along x   Difference over frames

$$I(x + u, y + v, t + 1) \approx I(x, y, t) + I_x \cdot u + I_y \cdot v + I_t$$

$$I(x + u, y + v, t + 1) - I(x, y, t) = +I_x \cdot u + I_y \cdot v + I_t$$

So: $I_x \cdot u + I_y \cdot v + I_t \approx 0 \quad \rightarrow \nabla I \cdot \begin{bmatrix} u & v \end{bmatrix}^{\mathrm{T}} + I_t = 0$

The brightness constancy constraint

Can we use this equation to recover image motion (u,v) at each pixel?
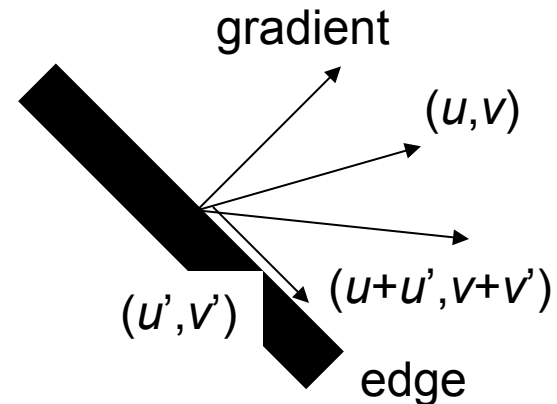
$$\nabla I \cdot \begin{bmatrix} u & v \end{bmatrix}^{\mathrm{T}} + I_t = 0$$

- How many equations and unknowns per pixel?

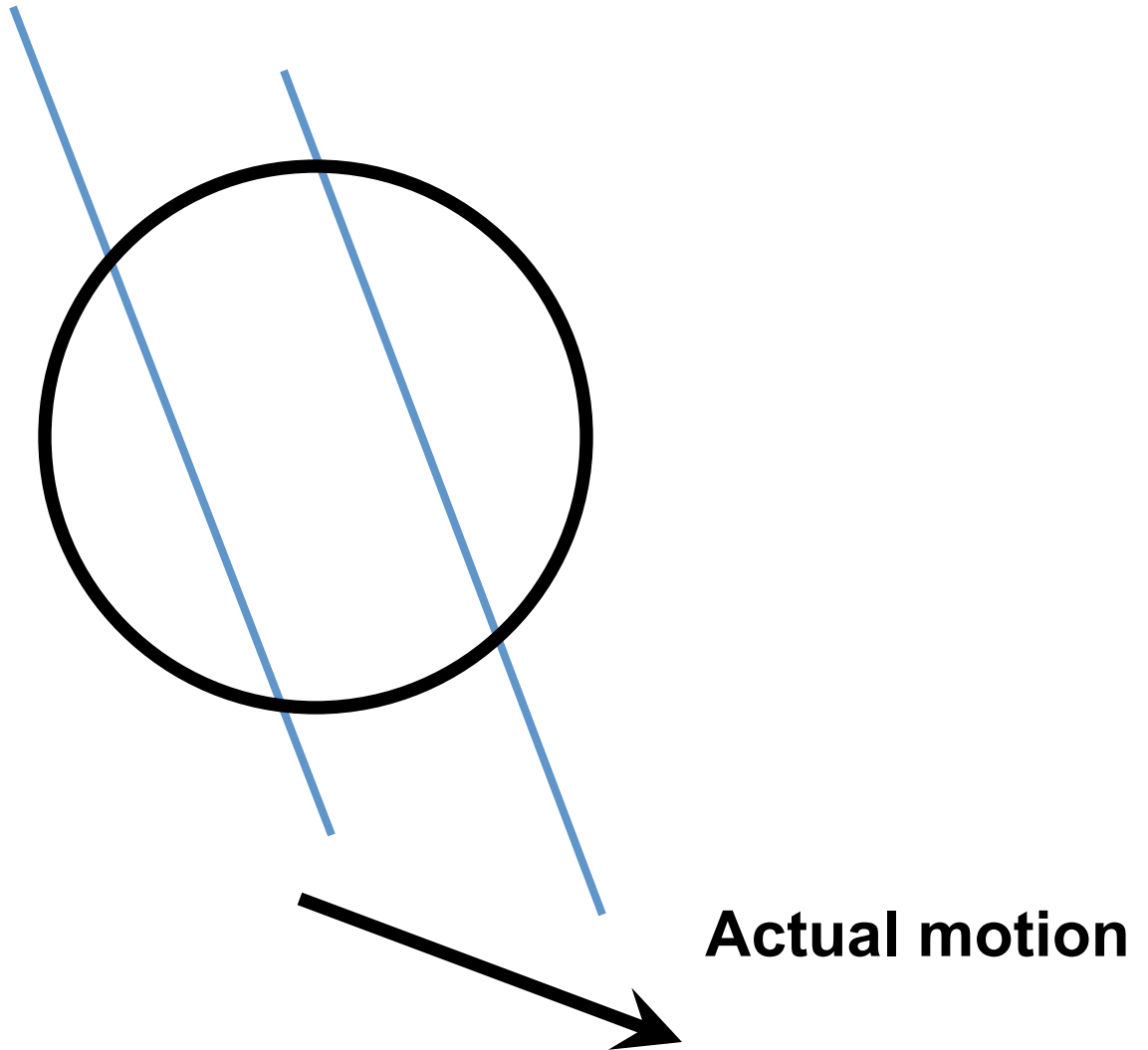 - One equation (this is a scalar equation!), two unknowns (u,v)

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

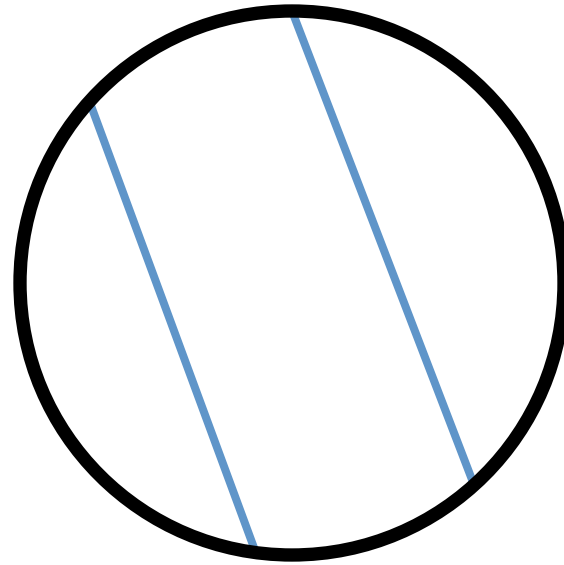If (u, v) satisfies the equation,
so does ($u+u'$, $v+v'$) if

$$\nabla I \cdot \begin{bmatrix} u' & v' \end{bmatrix}^{\mathrm{T}} = 0$$

gradient

(u,v)

$(u+u',v+v')$

(u',v')

edge

The aperture problem

Actual motion

The aperture problem

Perceived motion

The barber pole illusion

The barber pole illusion

Solving the ambiguity…

B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of th*
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:

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

25x2 2x1 25x1

Matching patches across images

- Overconstrained linear system

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

25x2   2x1   25x1

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

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

$\qquad\qquad A^T A \qquad\qquad\qquad\qquad\qquad A^T b$

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

Conditions for solvability

Optimal (u, v) satisfies Lucas-Kanade equation

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

$$A^T A \qquad\qquad\qquad\qquad A^T b$$

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

- **A^{T}A** should be invertible
- **A^{T}A** should not be too small due to noise
  - eigenvalues λ_1 and λ_2 of **A^{T}A** should not be too small
- **A^{T}A** 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

Aperture problem
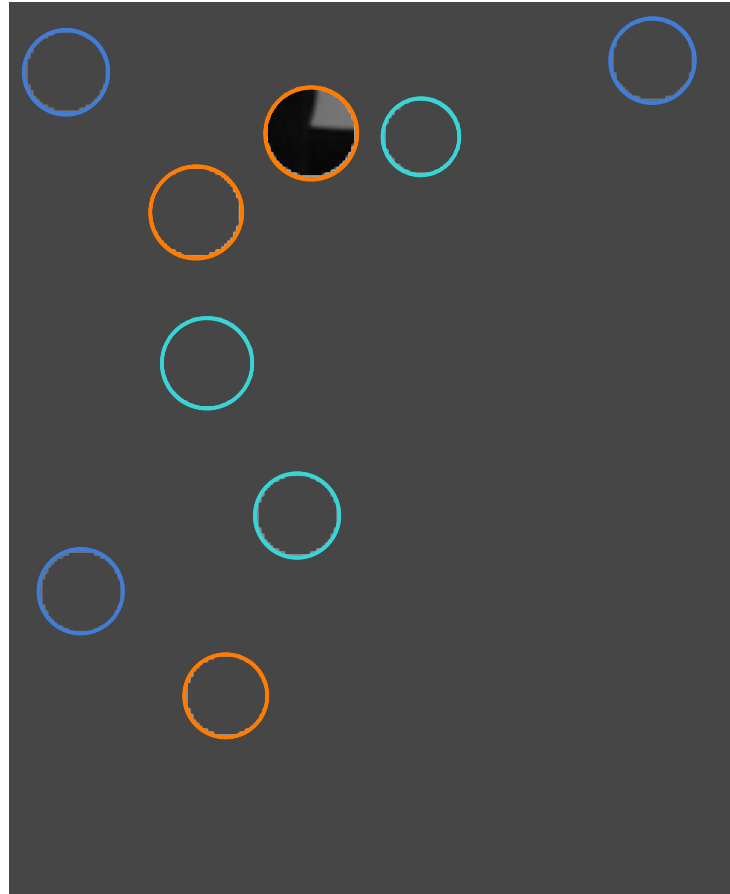
Edge
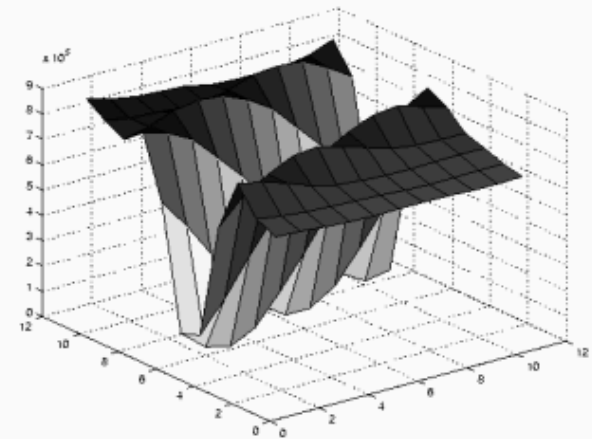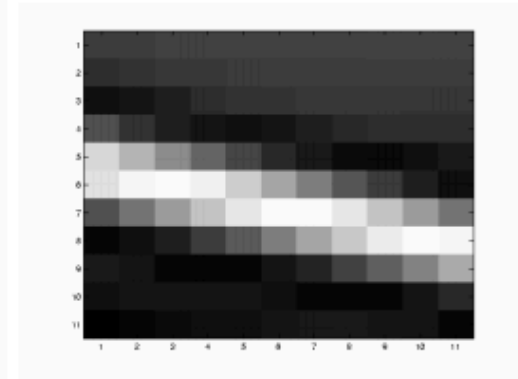
$$\sum \nabla I (\nabla I)^T$$

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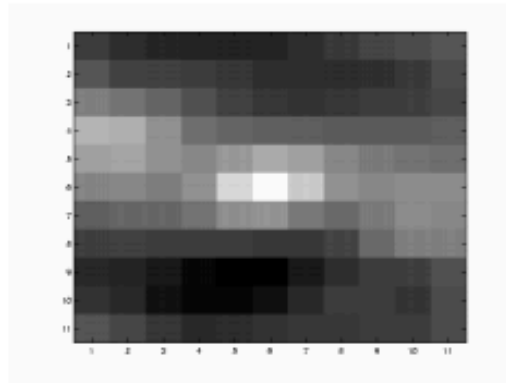 – large λ_1, small λ_2

Low Texture Region
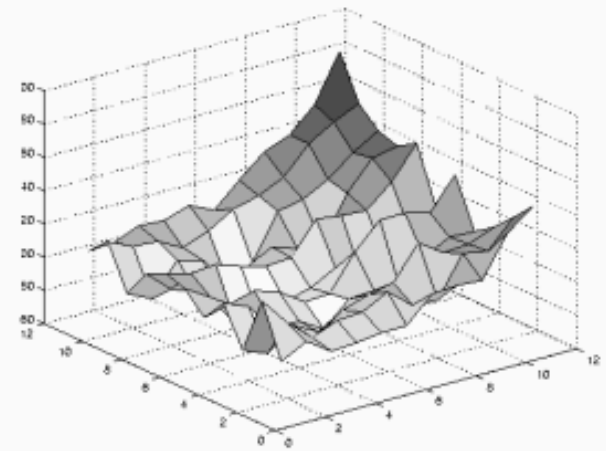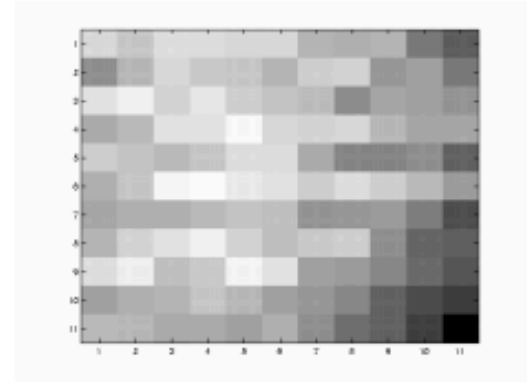
$$\sum \nabla I (\nabla I)^T$$

– gradients have small magnitude
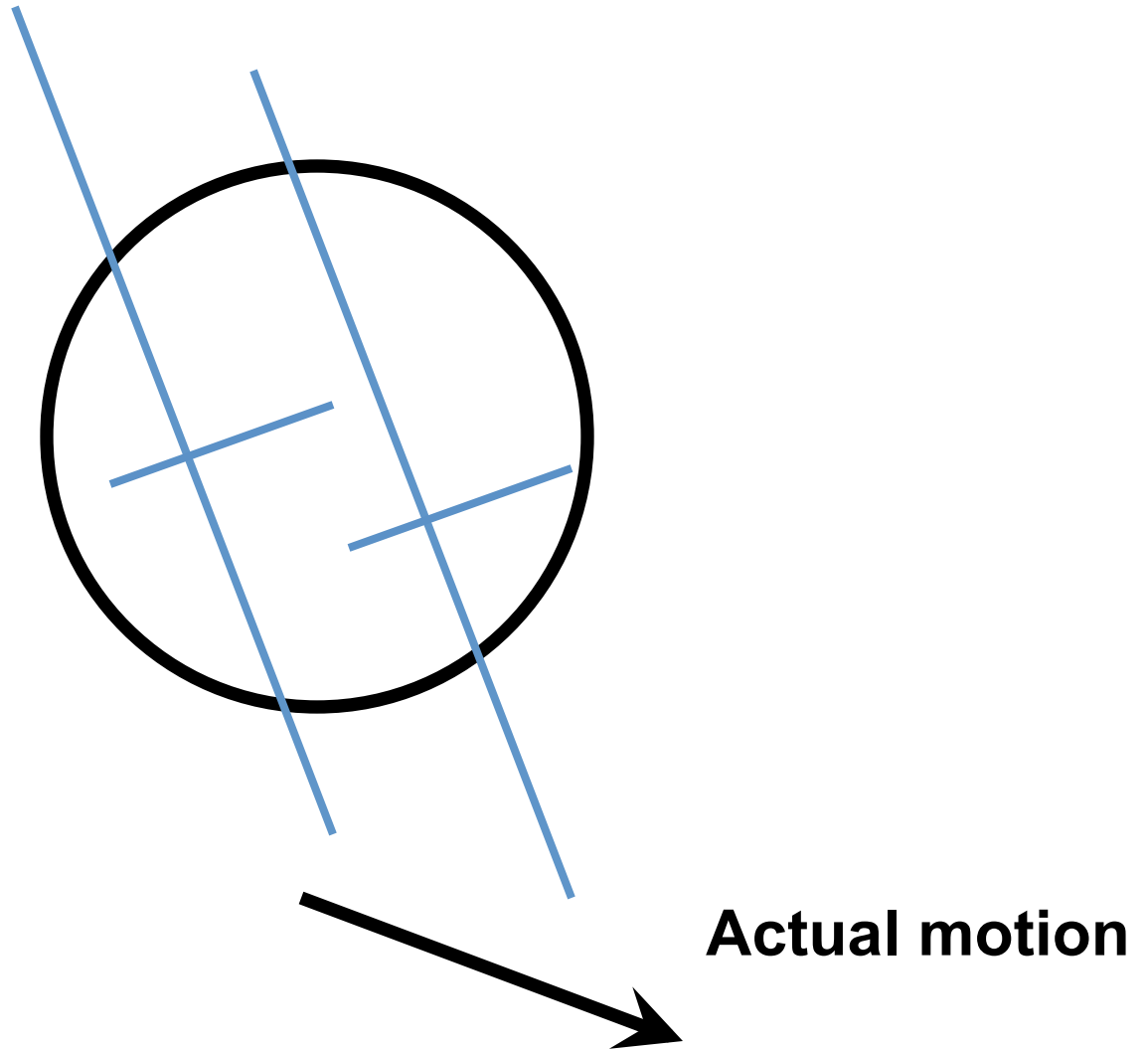– small λ_1, small λ_2

High Texture Region
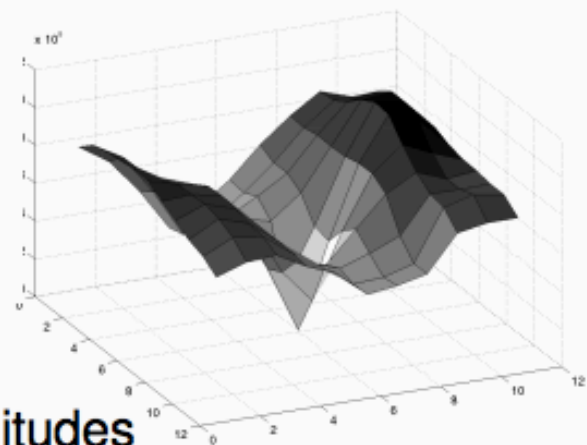
$$\sum \nabla I (\nabla I)^T$$

– gradients are different, large magnitudes
– large λ_1, large λ_2

The aperture problem resolved

Actual motion

The aperture problem resolved
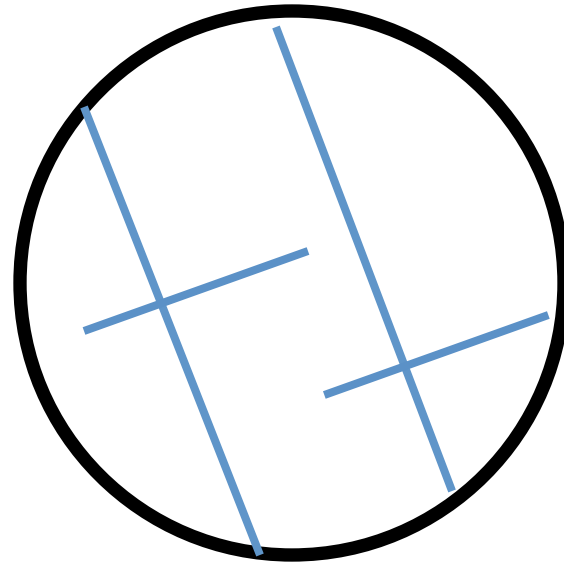
Perceived motion

Dealing with larger movements: Iterative refinement

Original (x,y) position

$$I_t = I(x', y', t+1) - I(x, y, t)$$

1. Initialize (x',y') = (x,y)

2. Compute (u,v) by

$$\left[\begin{array}{cc} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{array} \right] \left[\begin{array}{c} u \\ v \end{array} \right] = - \left[\begin{array}{c} \sum I_x I_t \\ \sum I_y I_t \end{array} \right]$$

2nd moment matrix for feature patch in first image

displacement

3. Shift window by (u, v): `x'=x'+u; y'=y'+v;`

4. Recalculate I_t

5. Repeat steps 2-4 until small change
   - Use interpolation for subpixel values
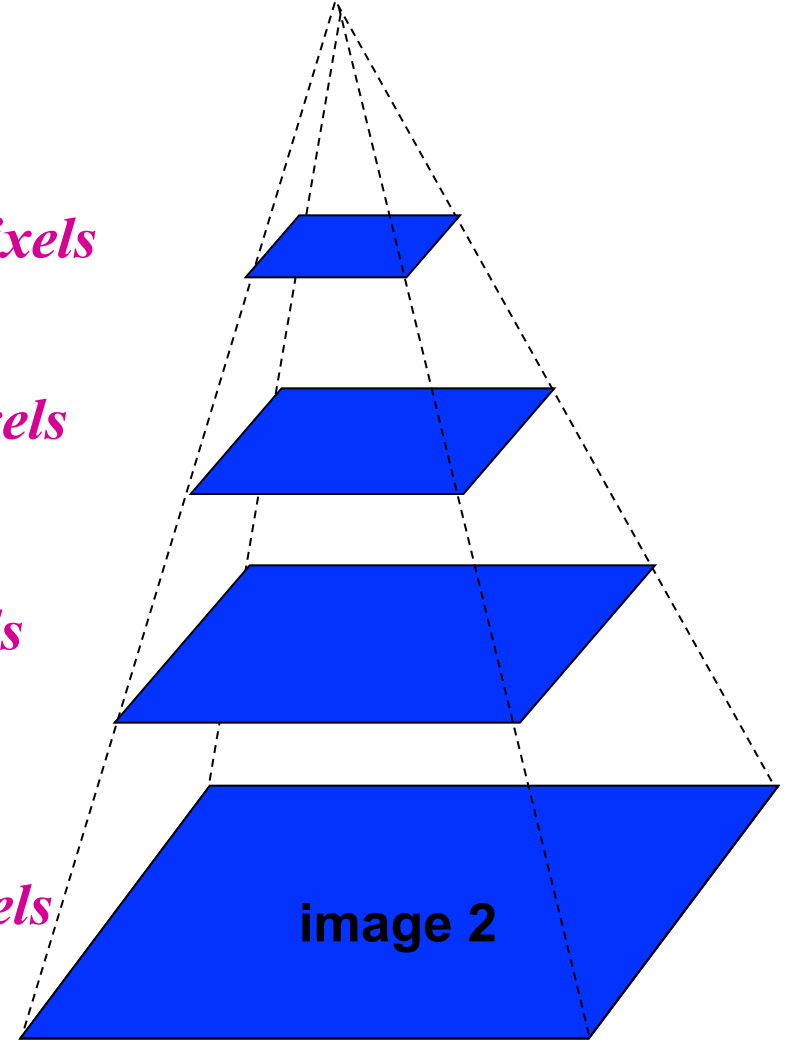
Coarse-to-fine optical flow estimation

run iterative L-K

warp & upsample

run iterative L-K

Gaussian pyramid of image 1 (t)
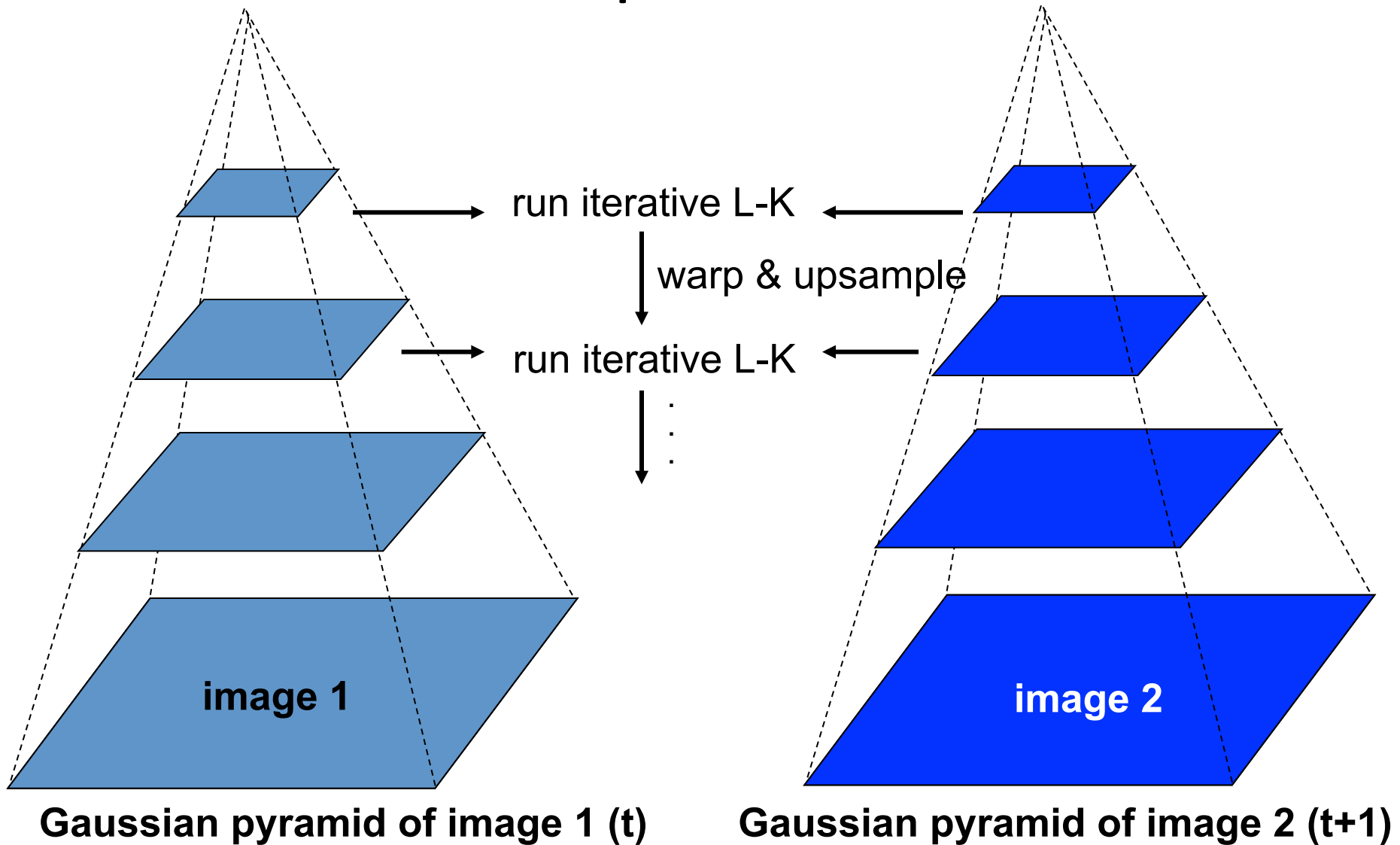
image 1

image 2

Gaussian pyramid of image 2 (t+1)

Coarse-to-fine optical flow estimation
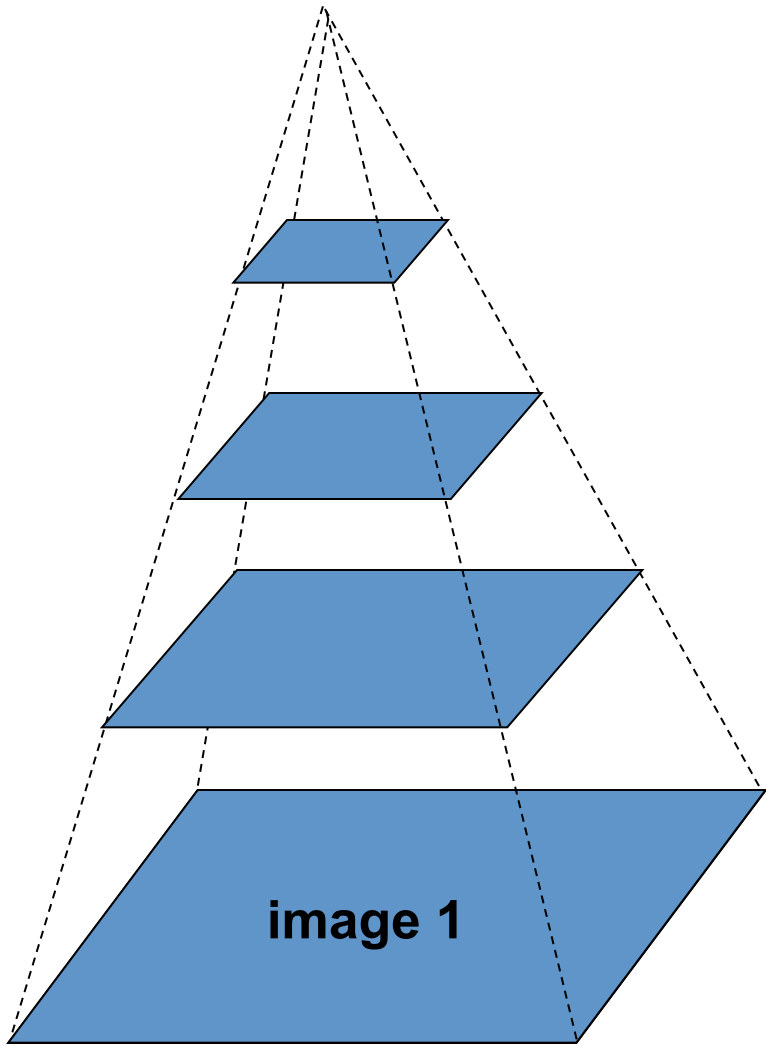
u=1.25 pixels

u=2.5 pixels

u=5 pixels

image 1

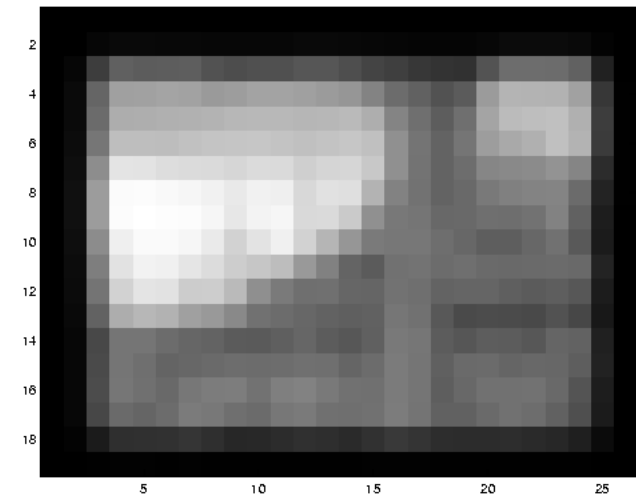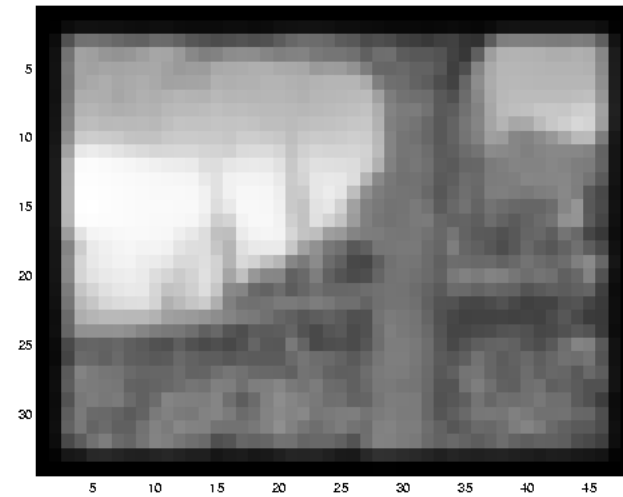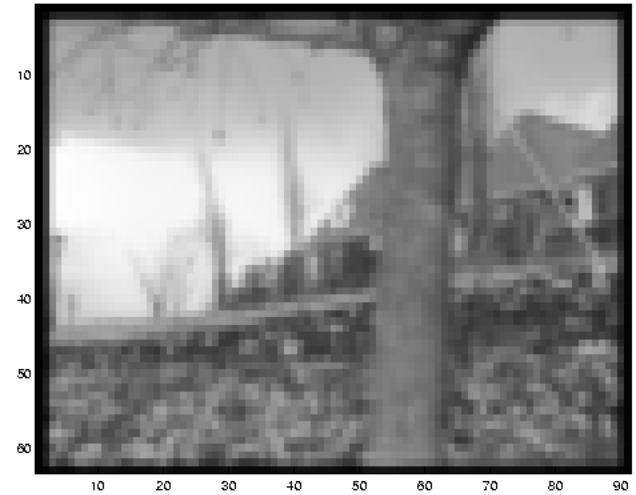*u=10 pixels*

image 2

Gaussian pyramid of image 1

Gaussian pyramid of image 2
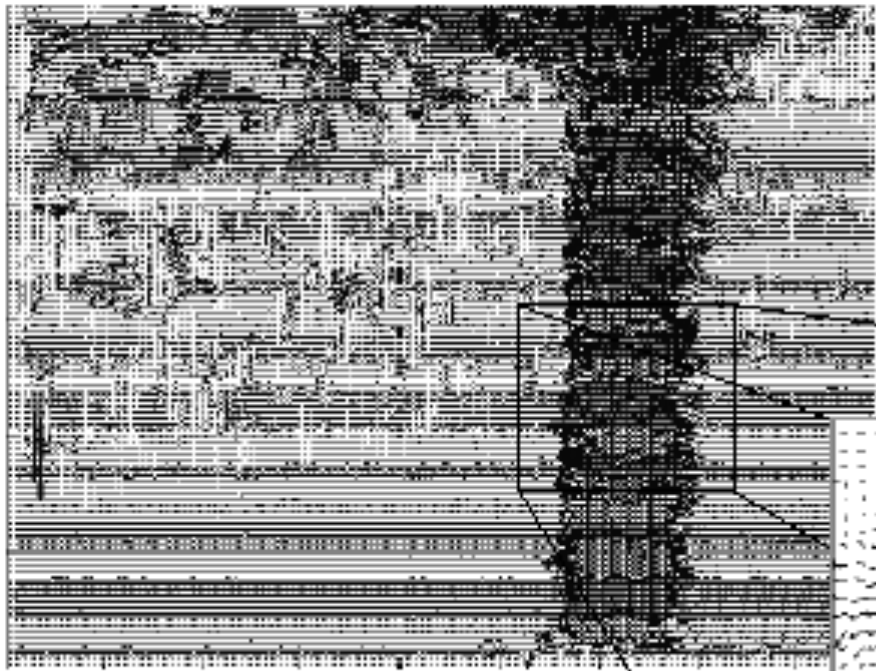
Example

Multi-resolution registration

Optical Flow Results
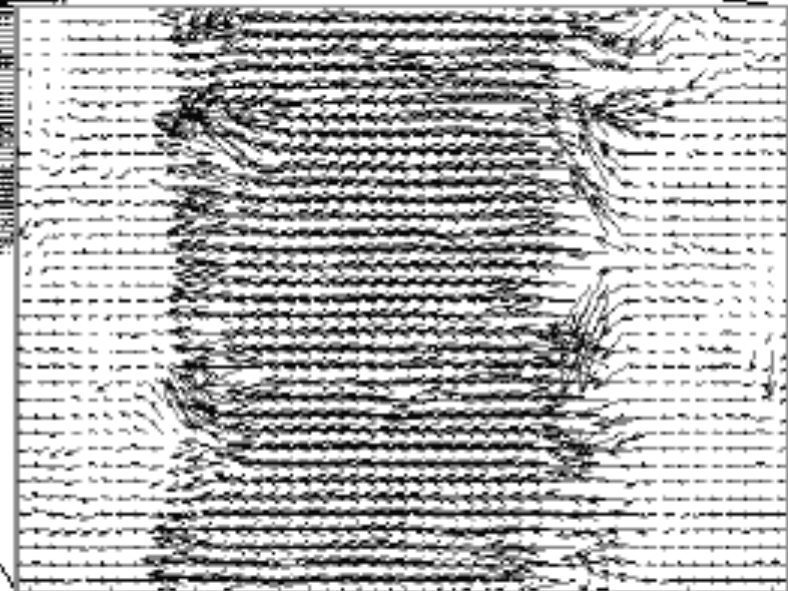

Lucas-Kanade
without pyramids

Fails in areas of large
motion
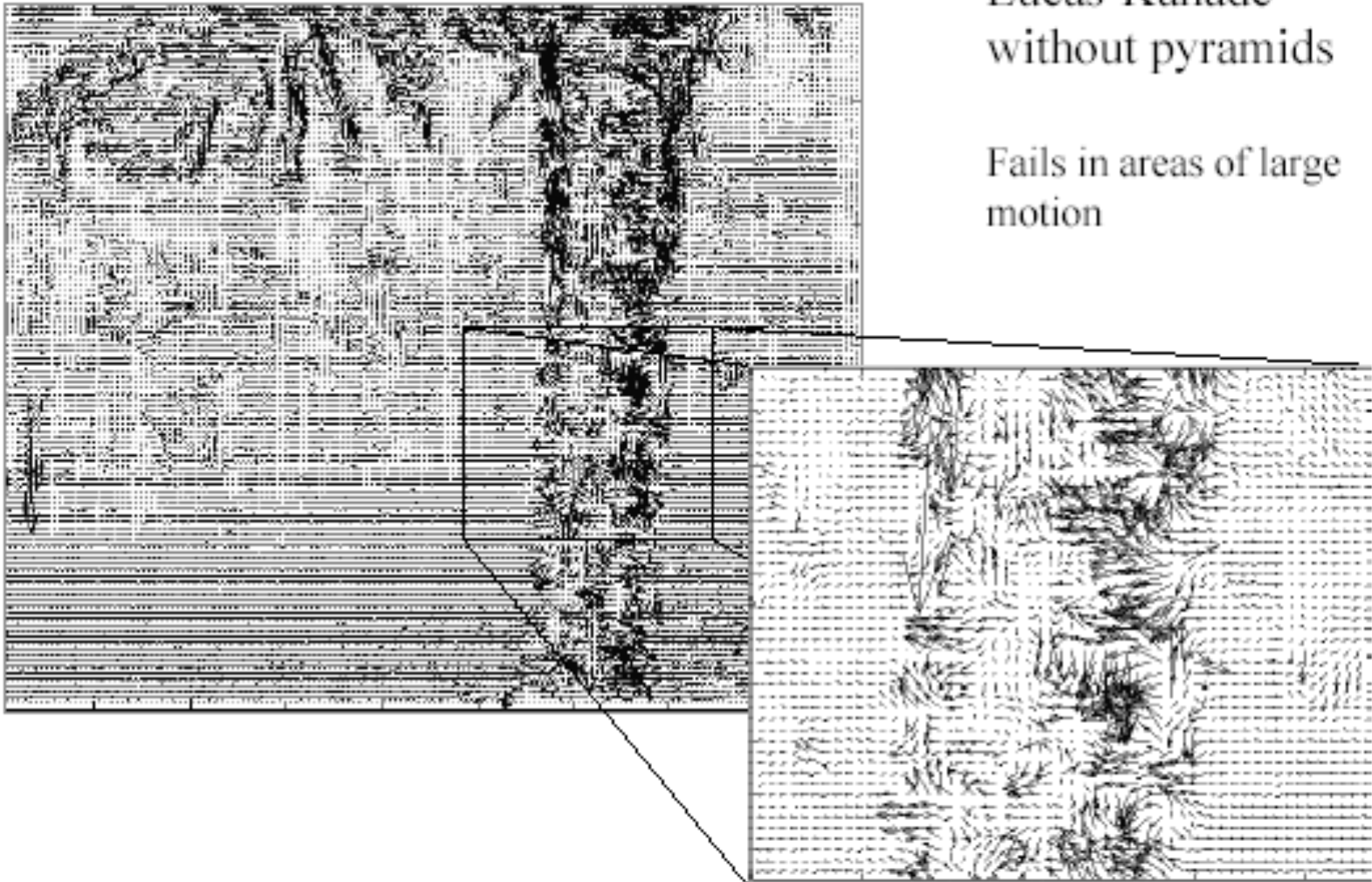
Optical Flow Results

Lucas-Kanade with Pyramids

Errors in Lucas-Kanade

- The motion is large
  - Possible Fix: Keypoint matching
- A point does not move like its neighbors
  - Possible Fix: Region-based matching
- Brightness constancy does not hold
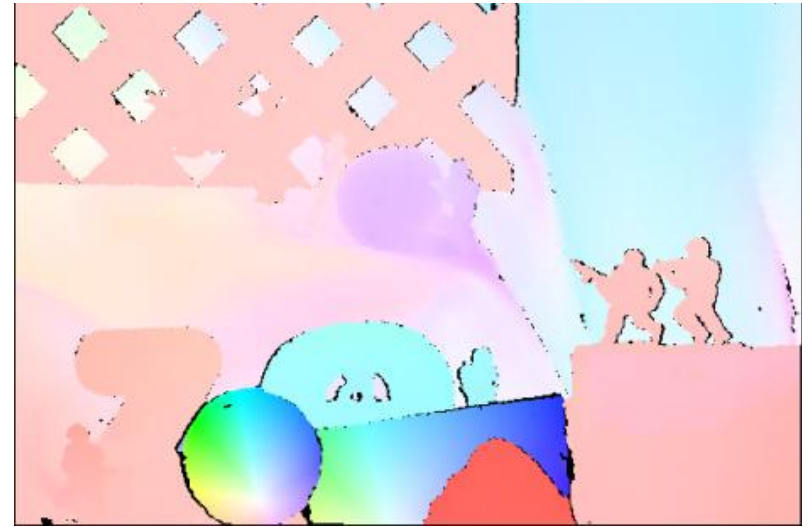 - Possible Fix: Gradient constancy
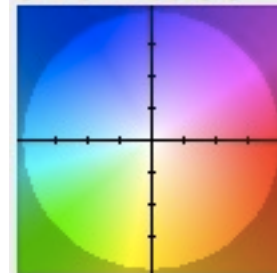
Flow quality evaluation

Flow quality evaluation
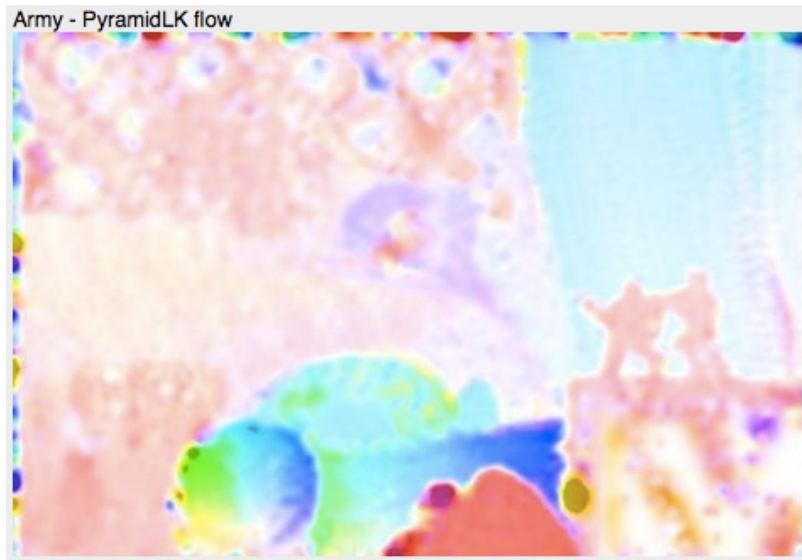
Flow quality evaluation

- Middlebury flow page
  - http://vision.middlebury.edu/flow/

Ground Truth

Color encoding of flow vectors

Flow quality evaluation

- Middlebury flow page
  - http://vision.middlebury.edu/flow/

Army - PyramidLK flow

Lucas-Kanade flow

Ground Truth
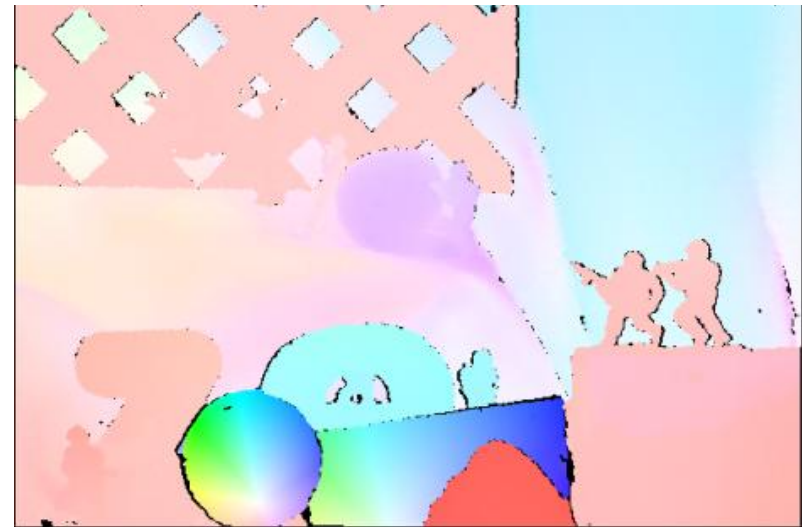
Color encoding of flow vectors

Flow quality evaluation
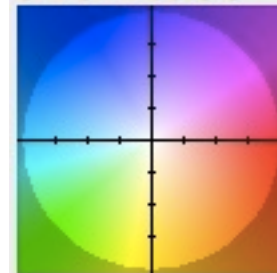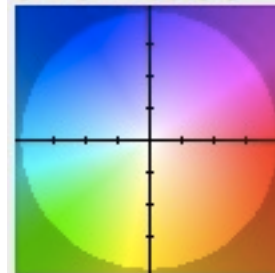
- Middlebury flow page
  - http://vision.middlebury.edu/flow/

Best-in-class alg
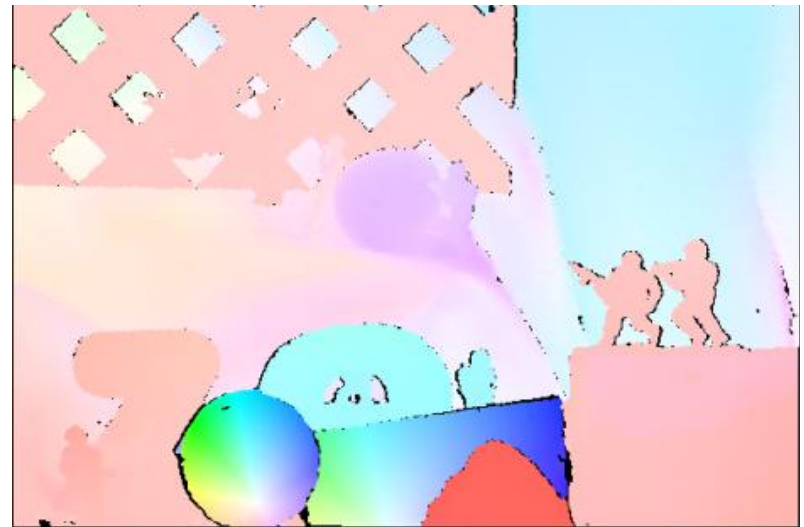
Ground Truth

Color encoding of flow vectors

Video stabilization

Video denoising

Video super resolution

Summary

- Major contributions from Lucas, Tomasi, Kanade
  - Tracking feature points
  - Optical flow
  - Stereo
  - Structure from motion

- Key ideas
  - By assuming brightness constancy, truncated Taylor expansion leads to simple and fast patch matching across frames
  - Coarse-to-fine registration