Lecture 13. Motion

Pattern Recognition and Computer Vision

Guanbin Li,
School of Computer Science and Engineering, Sun Yat-Sen University

扫码签到

What we will learn today?

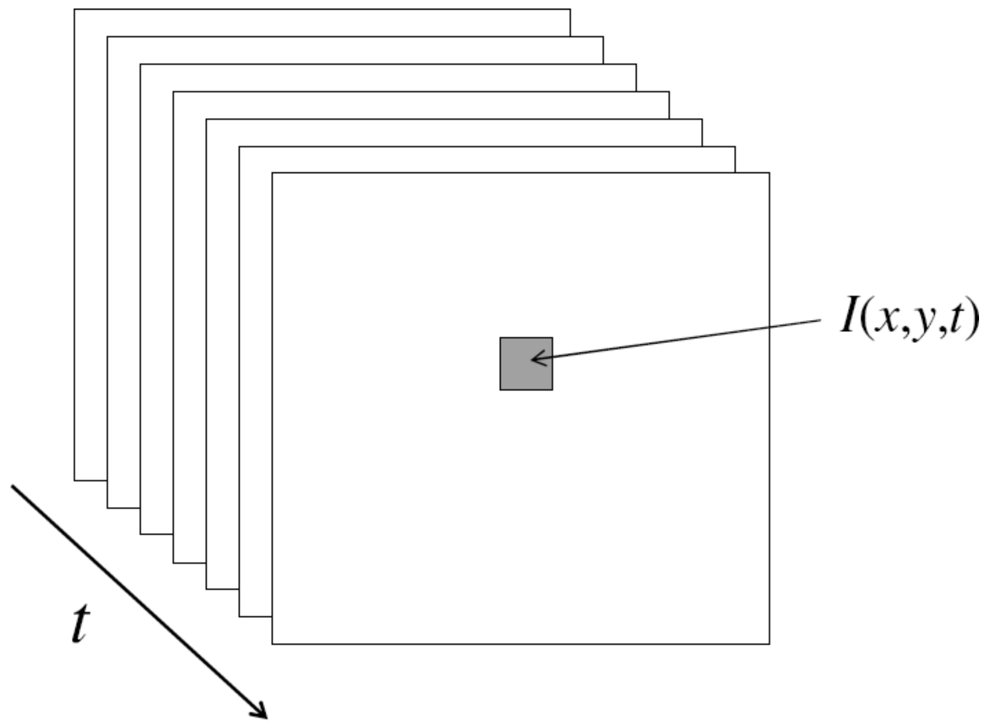
- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications

What we will learn today?

- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications

From images to videos

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

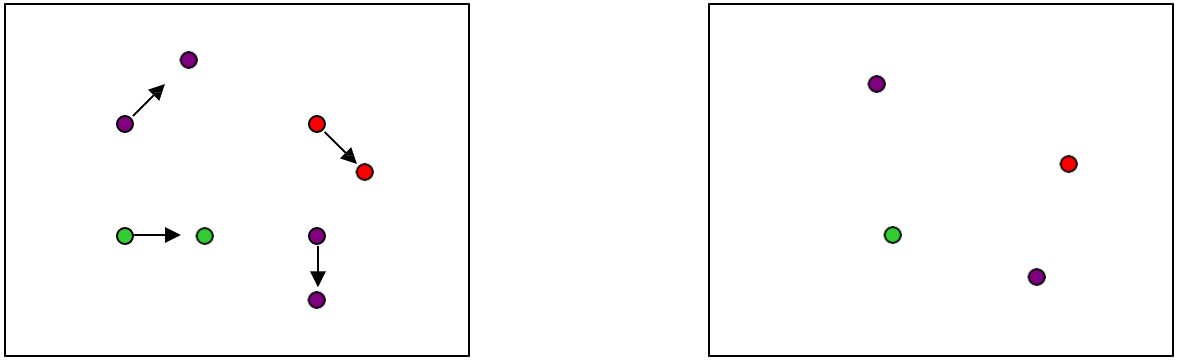


Optical flow

- Definition: optical flow is the *apparent* motion of brightness patterns in the image
- Note: apparent motion can be caused by lighting changes without any actual motion
  - – Like a stationary sphere under moving illumination

- **GOAL:** Recover image motion at each pixel from optical flow

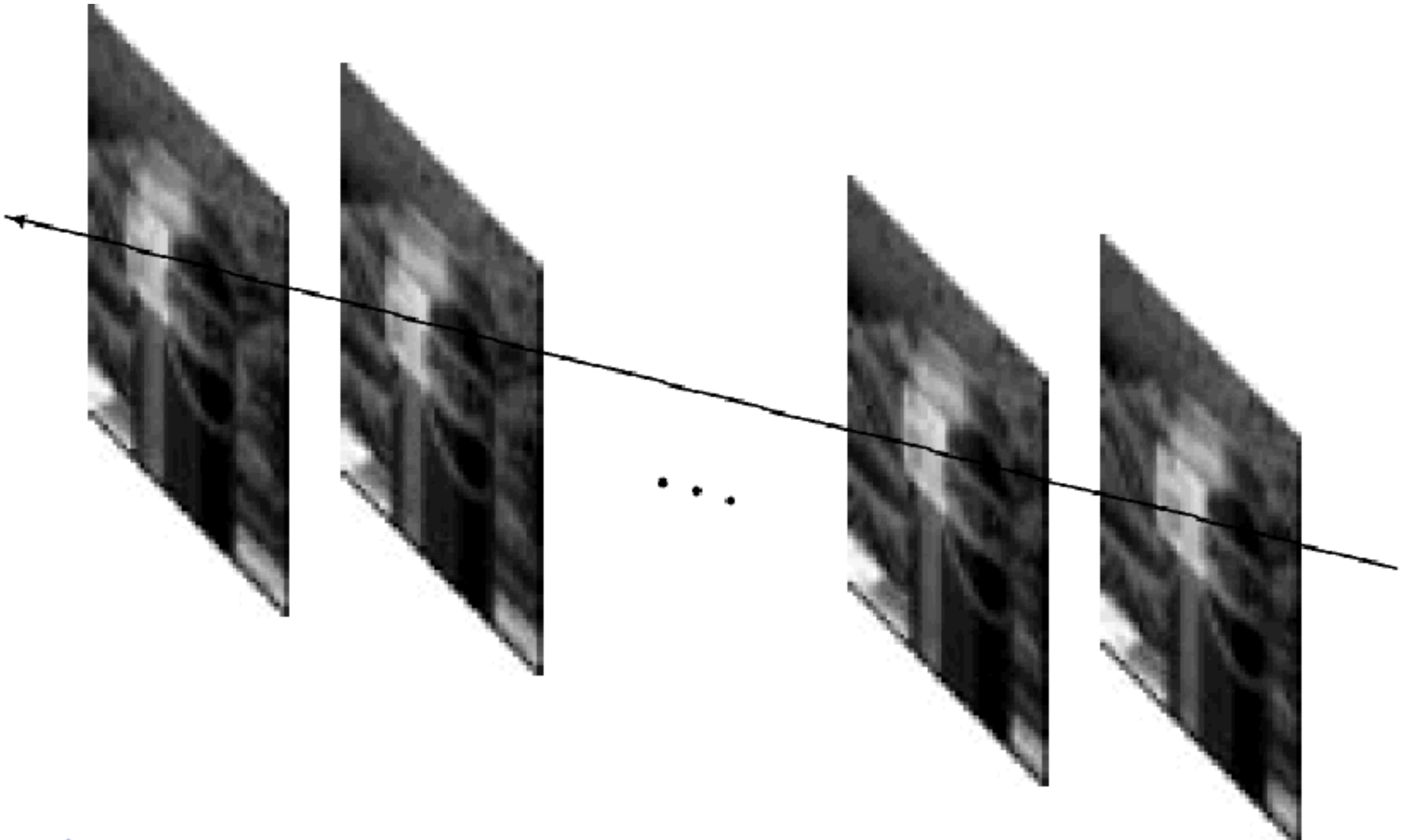
Estimating optical flow



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

- Given two subsequent frames, estimate the apparent motion field u(x,y), v(x,y) between them
- Key assumptions
  - **Small motion:** points do not move very far
  - **Spatial coherence:** points move like their neighbors
  - **Brightness constancy:** projection of the same point looks the same in every frame

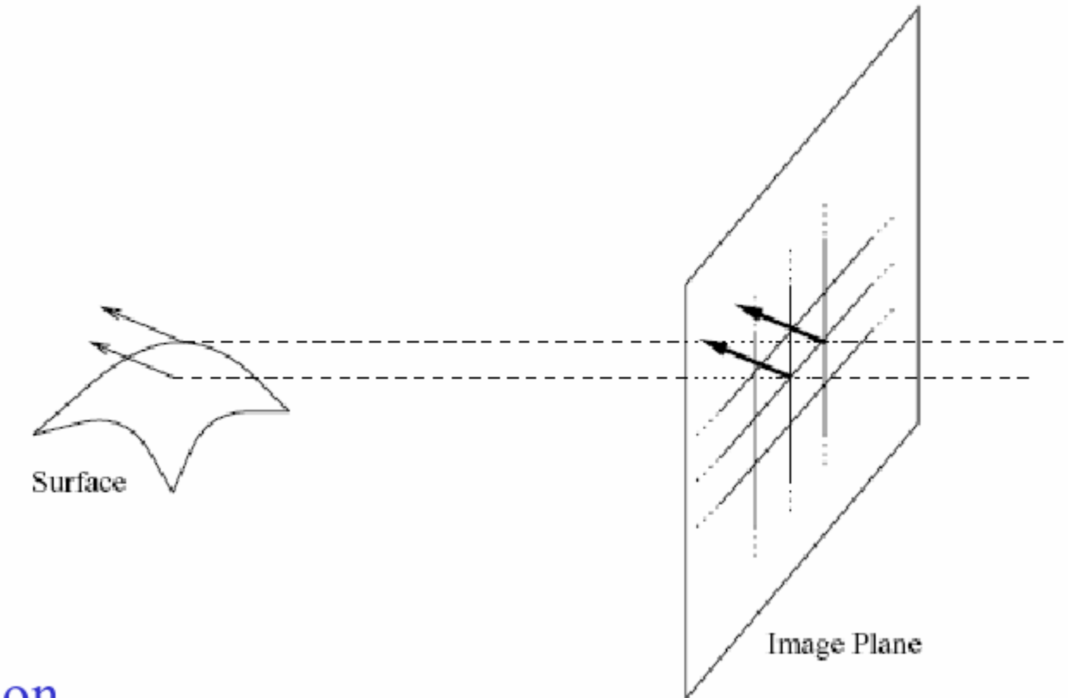
Key Assumptions: small motions



Assumption:

The image motion of a surface patch changes gradually over time.

Key Assumptions: spatial coherence



Assumption

* Neighboring points in the scene typically belong to the same surface and hence typically have similar motions.
* Since they also project to nearby points in the image, we expect spatial coherence in image flow.

Key Assumptions: brightness Constancy

0	0	0
0	100	0
0	0	0

t

0	0	0
100	0	0
0	0	0

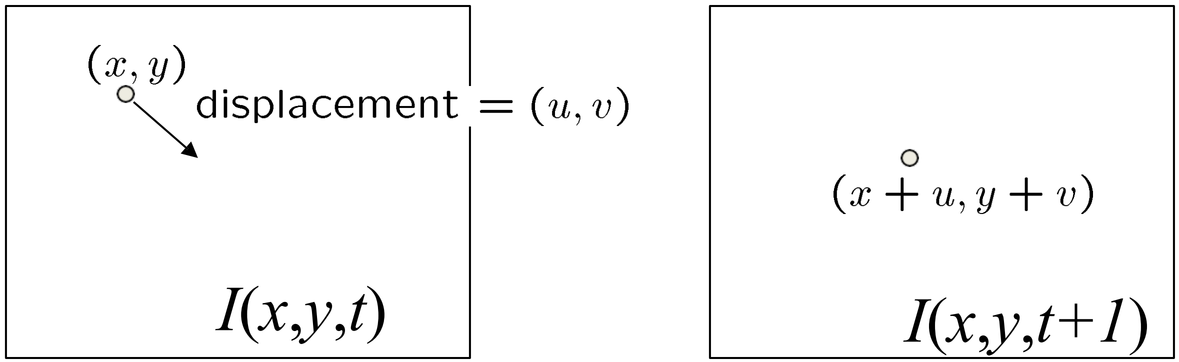
t+1

- Image measurements(e.g. brightness) in a small region remain the same although their location may change.

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

(assumption)

The brightness constancy constraint



- Brightness Constancy Equation:

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

Linearizing the right side using Taylor expansion:

Image derivative along x   Image derivative along t

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

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

Hence, $I_x u + I_y v + I_t \approx 0 \;\rightarrow\; \nabla I [u \; v]^T + I_t = 0$

Filters used to find the derivatives

$$\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \text{first image} \qquad \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \text{first image} \qquad \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \text{first image}$$

$$\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \text{second image} \qquad \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \text{second image} \qquad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{second image}$$

$$I_x \qquad I_y \qquad I_t$$

The brightness constancy constraint

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

$$\nabla I\,[u\ v]^T + I_t = 0$$

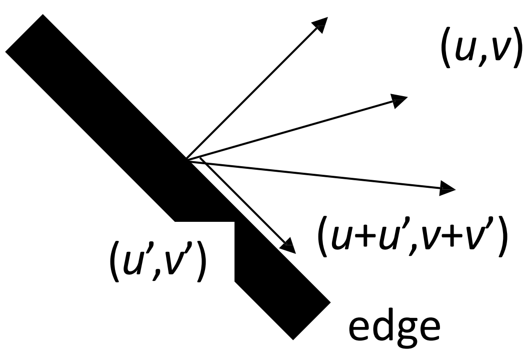
How many equations and unknowns per pixel?

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

The component of the flow 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

$$\nabla I \cdot [u'\ v']^T = 0$$



What we will learn today?

- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications

Solving the ambiguity…

$$\nabla I\,[u\ v]^T + I_t = 0$$

- 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

$$\begin{bmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{25}) & I_y(\mathbf{p}_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{25}) \end{bmatrix}$$

Lucas-Kanade flow

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

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?

- **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 anything to you?

Relate to Harris corner detector

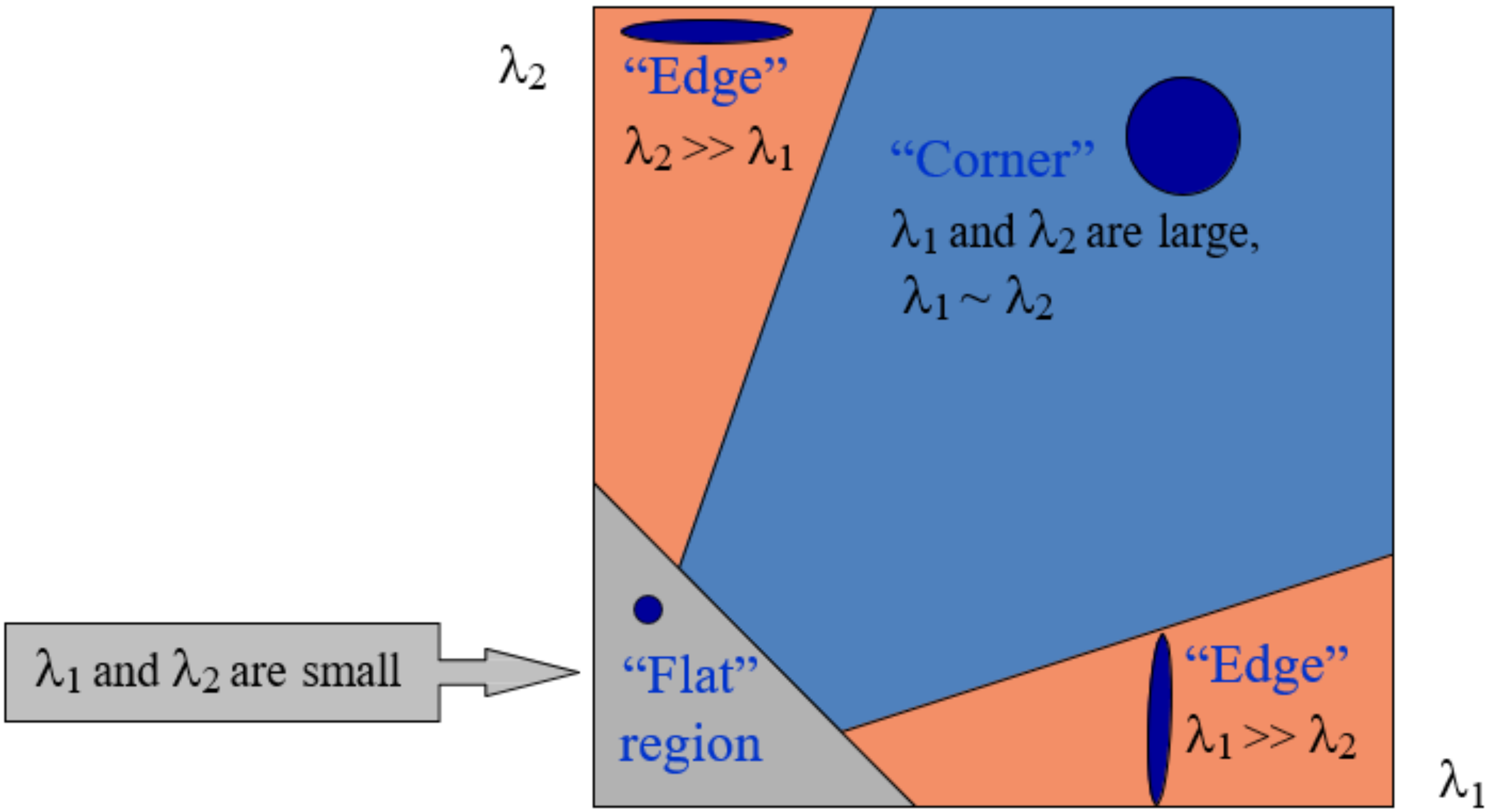
M = $A^T A$ is the second moment matrix !

(Harris corner detector…)

$$A^T A = \begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} = \sum \begin{bmatrix} I_x \\ I_y \end{bmatrix} [I_x \ I_y] = \sum \nabla I (\nabla I)^T$$

- Eigenvectors and eigenvalues of $A^T A$ relate to edge direction and magnitude
  - The eigenvector associated with the larger eigenvalue points in the direction of fastest intensity change
  - The other eigenvector is orthogonal to it

Interpreting the eigenvalues



Errors in Lukas-Kanade

What are the potential causes of errors in this procedure?

- Assumed A^TA is easily invertible
- Assumed there is not much noise in the image

- When our assumptions are violated
  - Brightness constancy is **not** satisfied
  - The motion is **not** small
  - A point does **not** move like its neighbors
    - window size is too large
    - what is the ideal window size?

When do the optical flow assumptions fail?

In other words, in what situations does the displacement of pixel patches not represent physical movement of points in space?

1. A uniform rotating sphere
   – nothing seems to move, yet it is rotating

2. Changing directions or intensities of lighting can make things seem to move
   – for example, if the specular highlight on a rotating sphere moves.

3. Muscle movement can make some spots on a cheetah move opposite direction of motion.
   – And infinitely more break downs of optical flow.

What we will learn today?

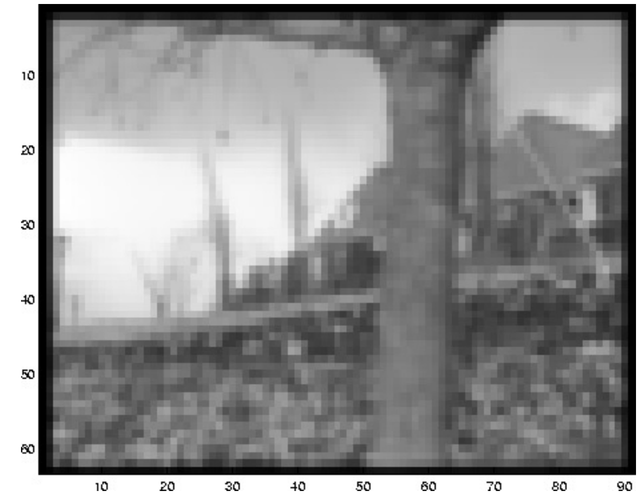
- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications

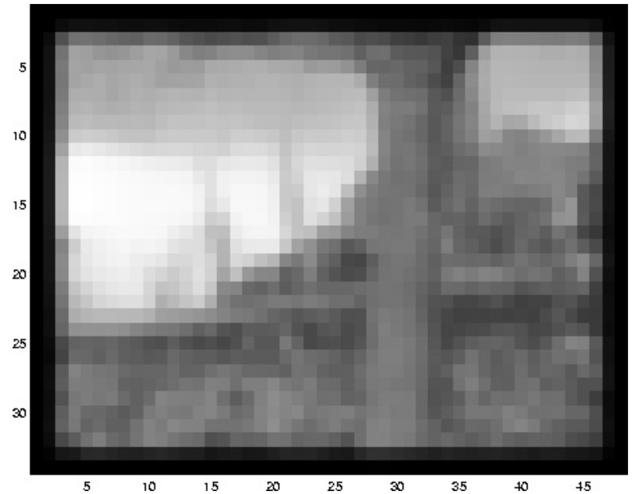
Recap

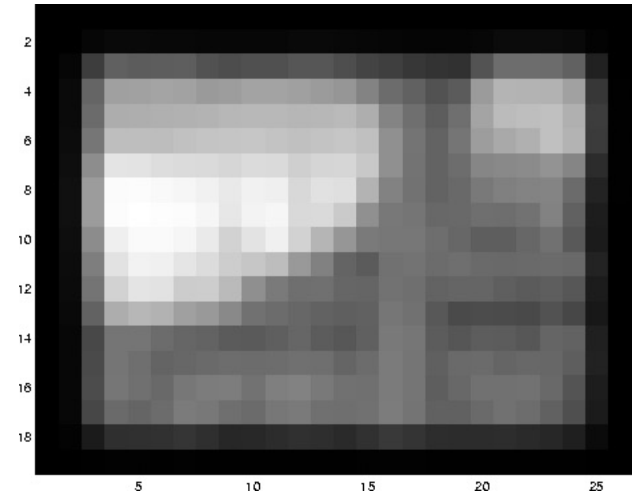
- Key assumptions (Errors in Lucas-Kanade)

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

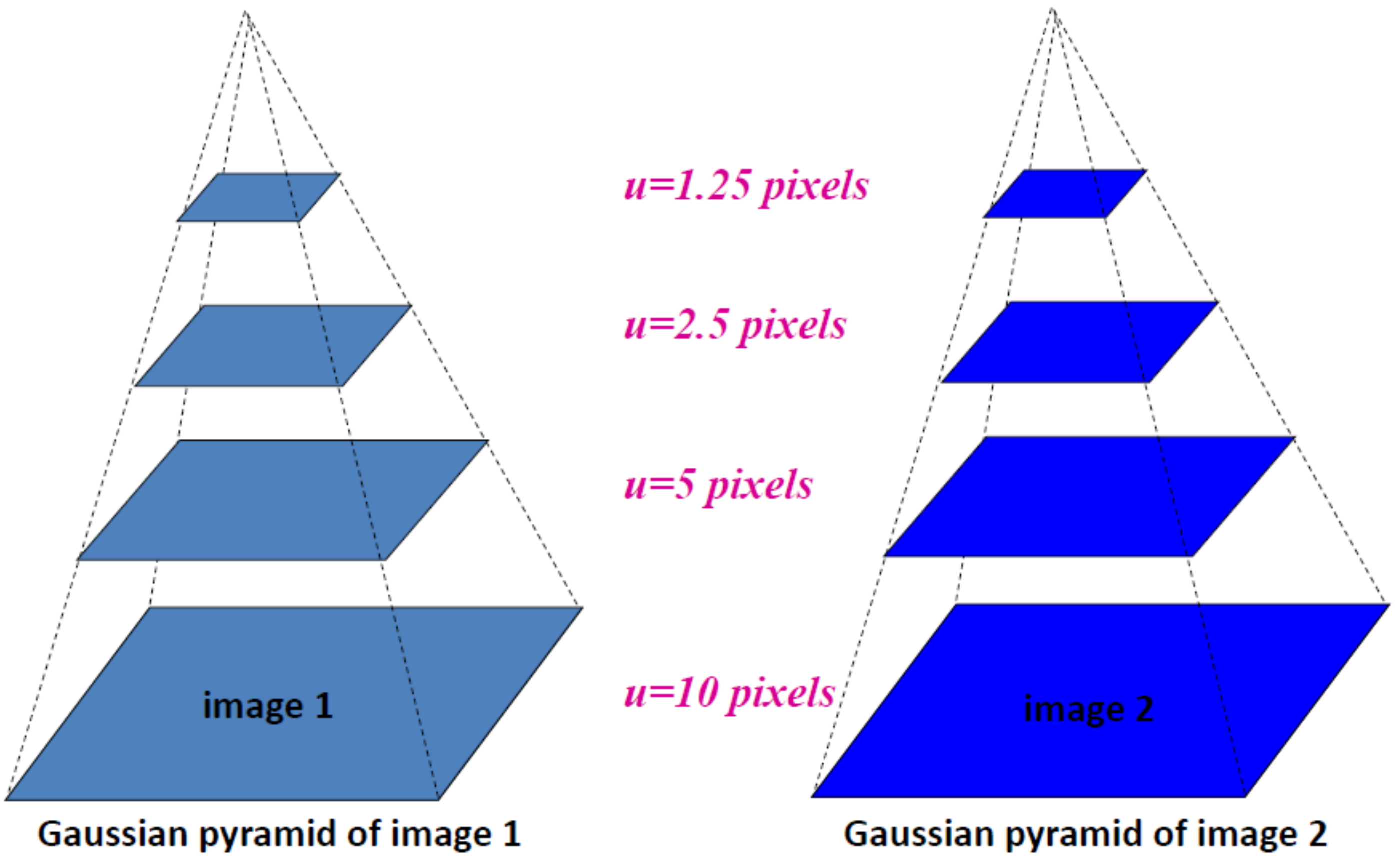
Reduce the resolution!



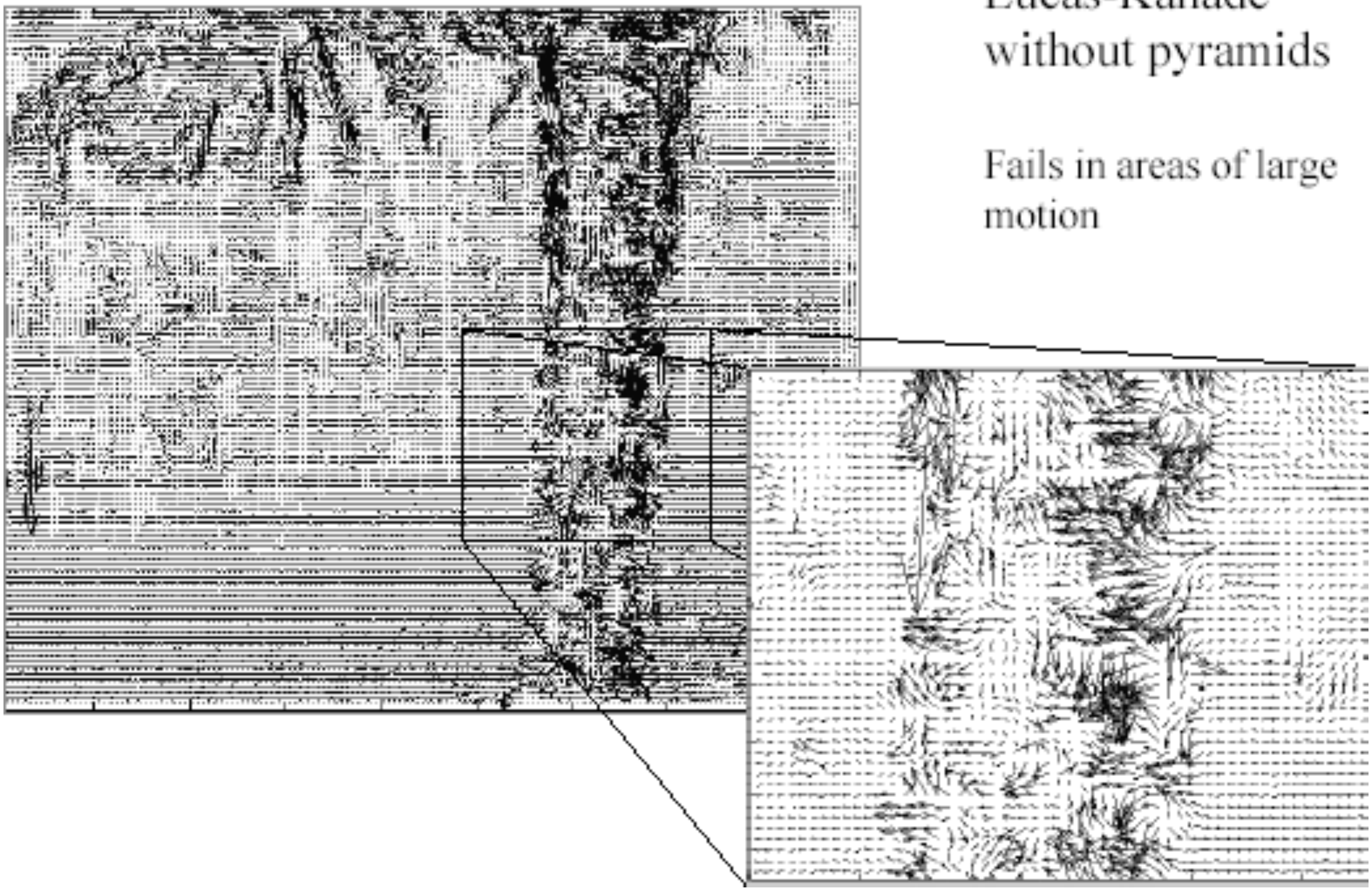




Coarse-to-fine optical flow estimation



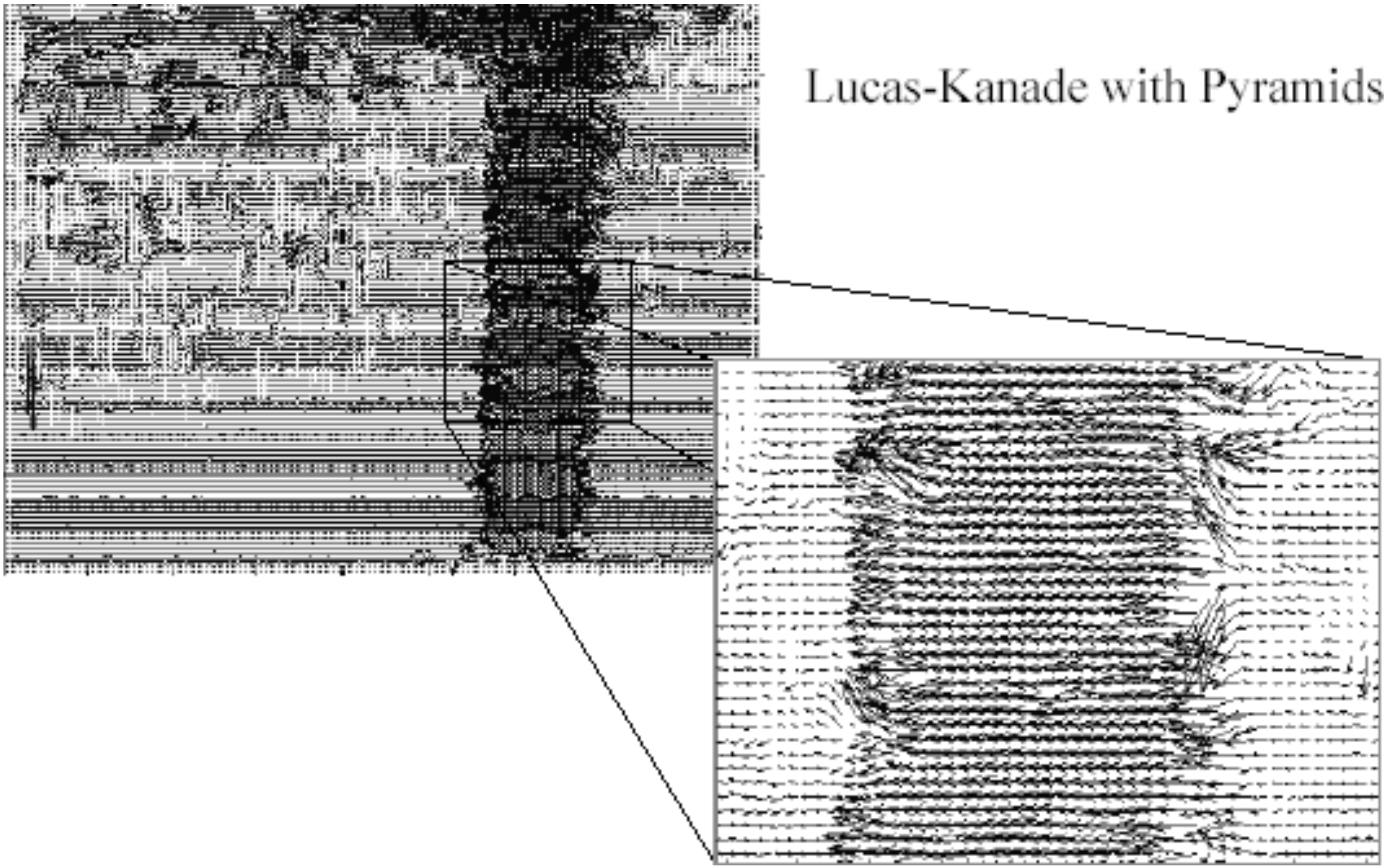
Optical Flow Results



Lucas-Kanade without pyramids

Fails in areas of large motion

Optical Flow Results



Lucas-Kanade with Pyramids

What we will learn today?

- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications

Horn-Schunk method for optical flow

- The flow is formulated as a global energy function which is should be minimized:

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

https://blog.csdn.net/u012841922/article/details/85273852

Horn-Schunk method for optical flow

- The flow is formulated as a global energy function which is should be minimized:
- The first part of the function is the brightness consistency.

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

Horn-Schunk method for optical flow

- The flow is formulated as a global energy function which is should be minimized:
- The second part is the smoothness constraint. It’s trying to make sure that the changes between pixels are small.

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

Horn-Schunk method for optical flow

- The flow is formulated as a global energy function which is should be minimized:

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

- This functional can be minimized by solving the associated multi-dimensional Euler–Lagrange equations giving

$$\frac{aL}{au}-\frac{a}{ax}\frac{aL}{au_x}-\frac{a}{ay}\frac{aL}{au_y}=0$$

$$\frac{aL}{a\mathrm{v}}-\frac{a}{ax}\frac{aL}{av_x}-\frac{a}{ay}\frac{aL}{a\mathrm{v}_y}=0$$

$$I_x(I_x u+I_y v+I_t)-\alpha^2\Delta u=0$$
$$I_y(I_x u+I_y v+I_t)-\alpha^2\Delta v=0$$

Horn-Schunk method for optical flow

$$I_x(I_x u + I_y v + I_t) - \alpha^2 \Delta u = 0$$
$$I_y(I_x u + I_y v + I_t) - \alpha^2 \Delta v = 0$$

$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is called the Lagrange operator. In practice, it is measured

$$\Delta u(x,y) = \overline{u}(x,y) - u(x,y)$$

$$\bar{u}_{i,j,k} = \frac{1}{6}\{u_{i-1,j,k} + u_{i,j+1,k} + u_{i+1,j,k} + u_{i,j-1,k}\} + \frac{1}{12}\{u_{i-1,j-1,k} + u_{i-1,j+1,k} + u_{i+1,j+1,k} + u_{i+1,j-1,k}\},$$
$$\bar{v}_{i,j,k} = \frac{1}{6}\{v_{i-1,j,k} + v_{i,j+1,k} + v_{i+1,j,k} + v_{i,j-1,k}\} + \frac{1}{12}\{v_{i-1,j-1,k} + v_{i-1,j+1,k} + v_{i+1,j+1,k} + v_{i+1,j-1,k}\},$$

- where $\overline{u}(x,y)$ is the weighted average of u measured at (x,y).

1/12	1/6	1/12
1/6	-1 i,j,k	1/6
1/12	1/6	1/12

$$(I_x^2 + \alpha^2)u + I_x I_y v = \alpha^2 \overline{u} - I_x I_t$$
$$I_x I_y u + (I_y^2 + \alpha^2)v = \alpha^2 \overline{v} - I_y I_t$$

- Which is linear in u and v and can be solved analytically for each pixel individually.

Iterative Horn-Schunk

- But since the solution depends on the neighboring values of the flow field, it must be repeated once the neighbors have been updated.
- So instead, we can iteratively solve for u and v using:

$$u^{k+1} = \overline{u}^k - \frac{I_x(I_x\overline{u}^k + I_y\overline{v}^k + I_t)}{\alpha^2 + I_x^2 + I_y^2}$$

$$v^{k+1} = \overline{v}^k - \frac{I_y(I_x\overline{u}^k + I_y\overline{v}^k + I_t)}{\alpha^2 + I_x^2 + I_y^2}$$

$\overline{u}^k$ and $\overline{v}^k$ is initially set to 0, and iterate until

$$|\lambda_k - \lambda_{k-1}| < \text{thresh}$$

$$\lambda = I_x u + I_y v + I_t$$

What we will learn today?

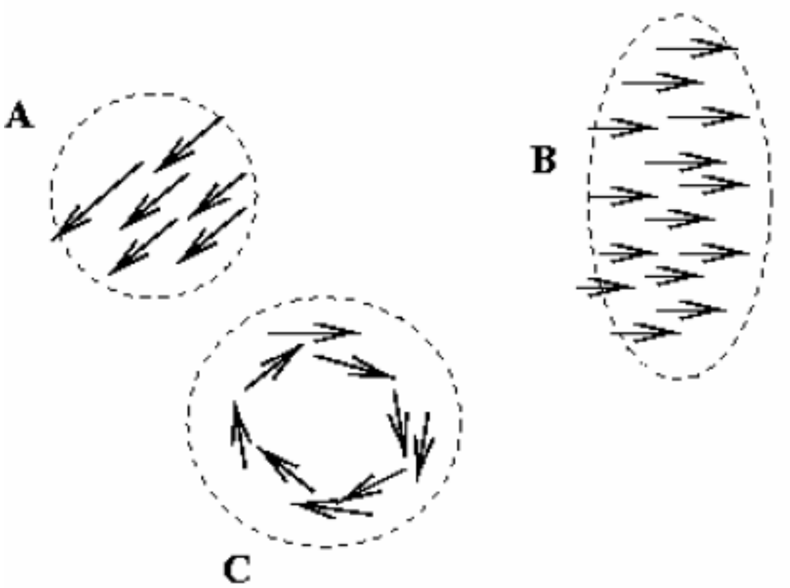
- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications

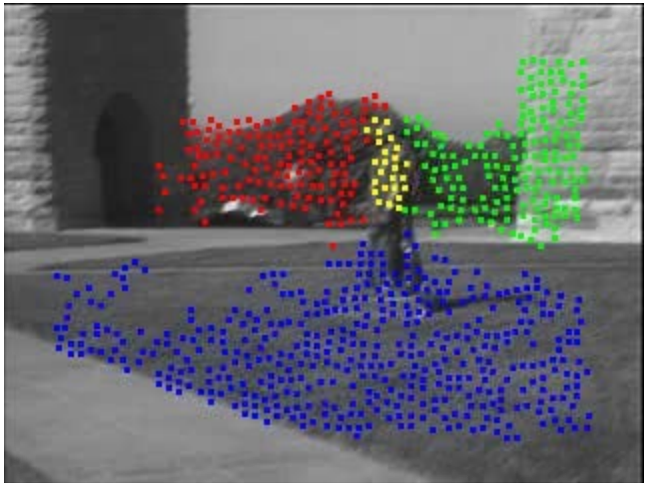
Uses of motion

- Tracking features
- Segmenting objects based on motion cues
- Learning dynamical models
- Improving video quality
  - Motion stabilization
  - Super resolution
- Tracking objects
- Recognizing events and activities

Segmenting objects based on motion cues

- Motion segmentation
  - Segment the video into multiple *coherently* moving objects

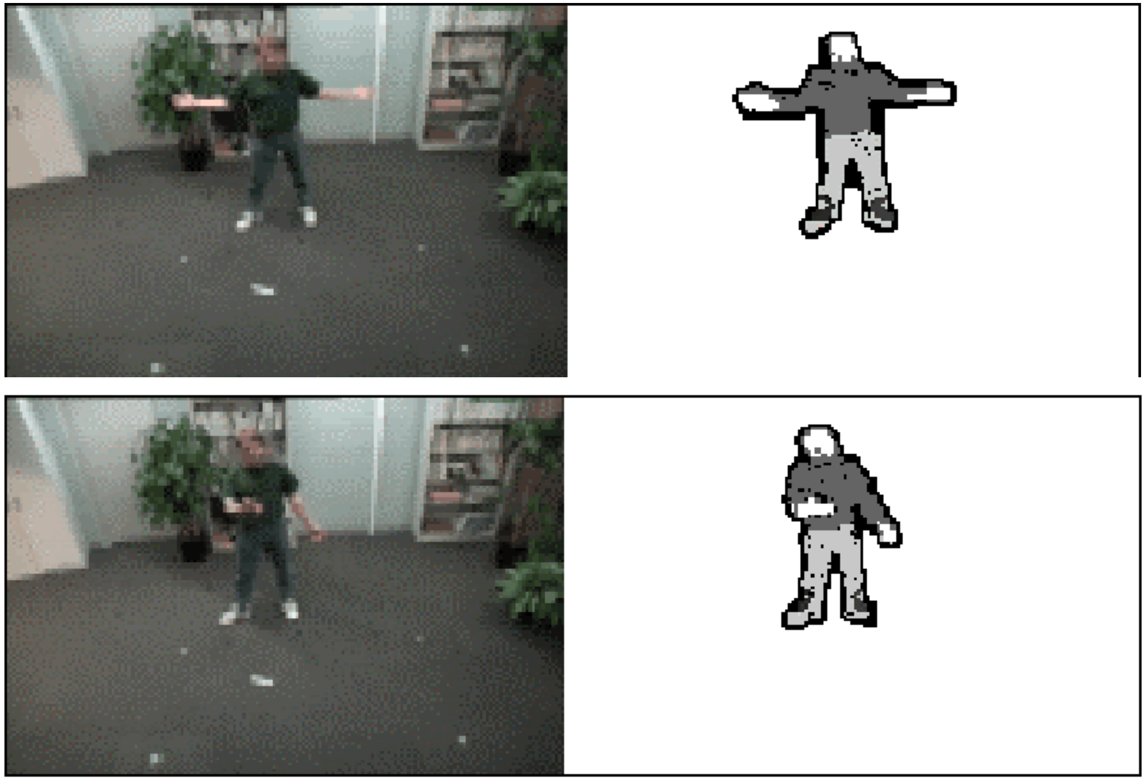




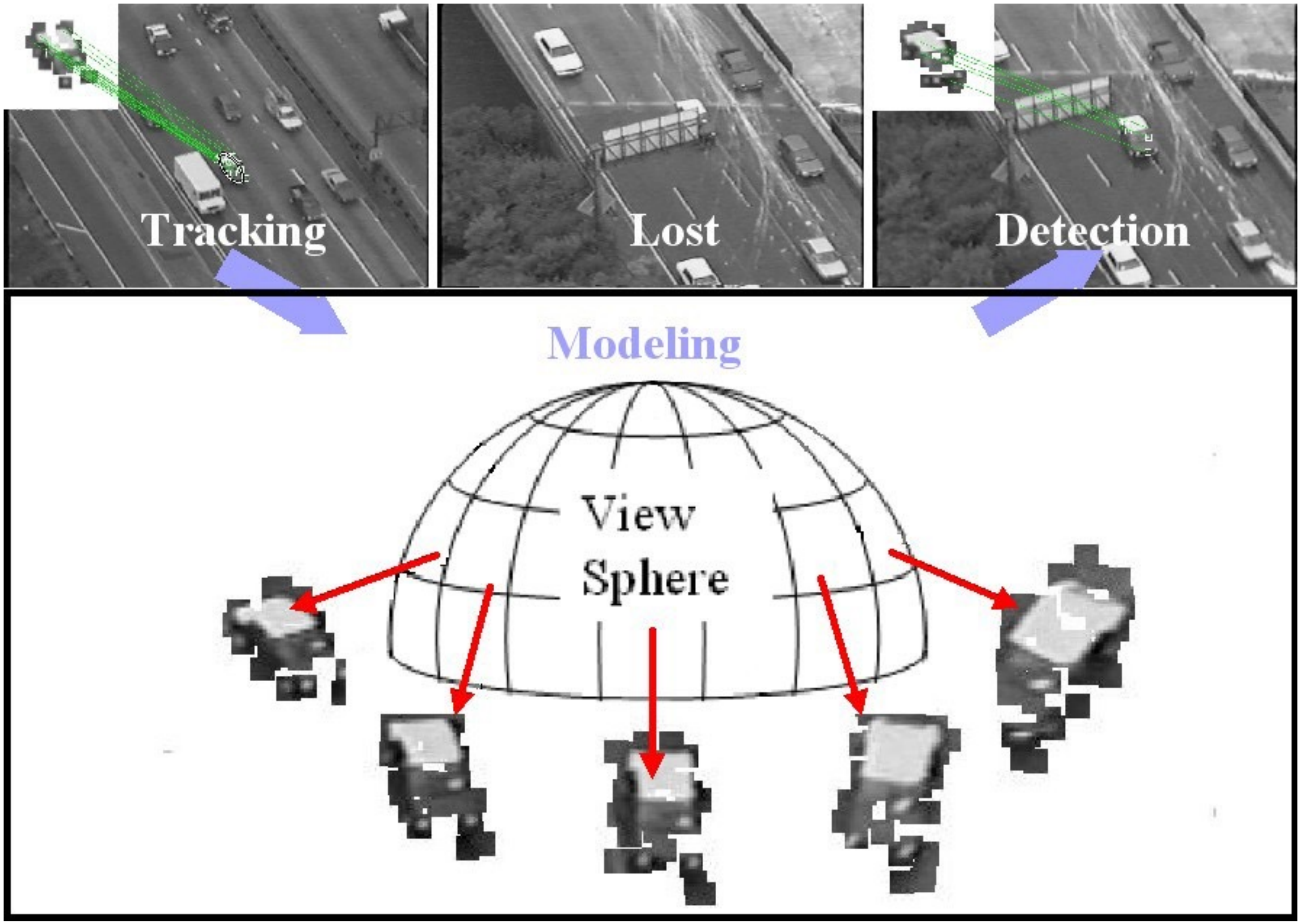
S. J. Pundlik and S. T. Birchfield, Motion Segmentation at Any Speed, Proceedings of the British Machine Vision Conference (BMVC) 2006

Segmenting objects based on motion cues

- Background subtraction
  - A static camera is observing a scene
  - Goal: separate the static *background* from the moving *foreground*



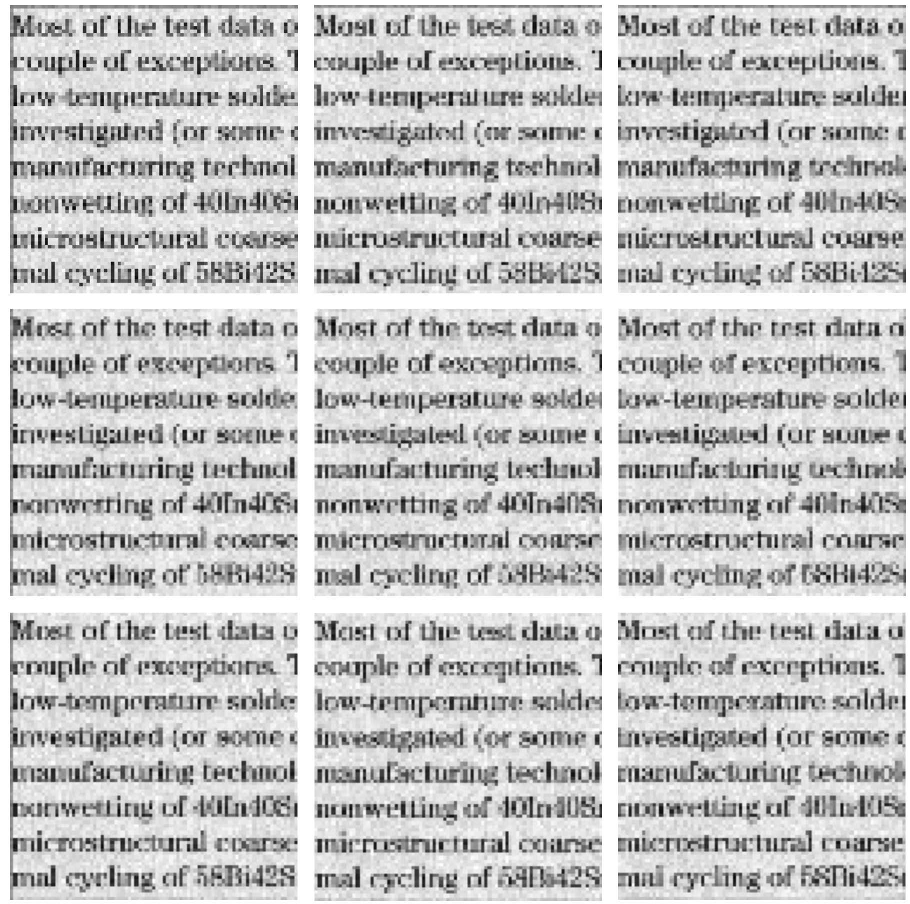
Tracking objects



Z.Yin and R.Collins, "On-the-fly Object Modeling while Tracking," *IEEE Computer Vision and Pattern Recognition (CVPR '07),* Minneapolis, MN, June 2007.

Super-resolution

Example: A set of low quality images

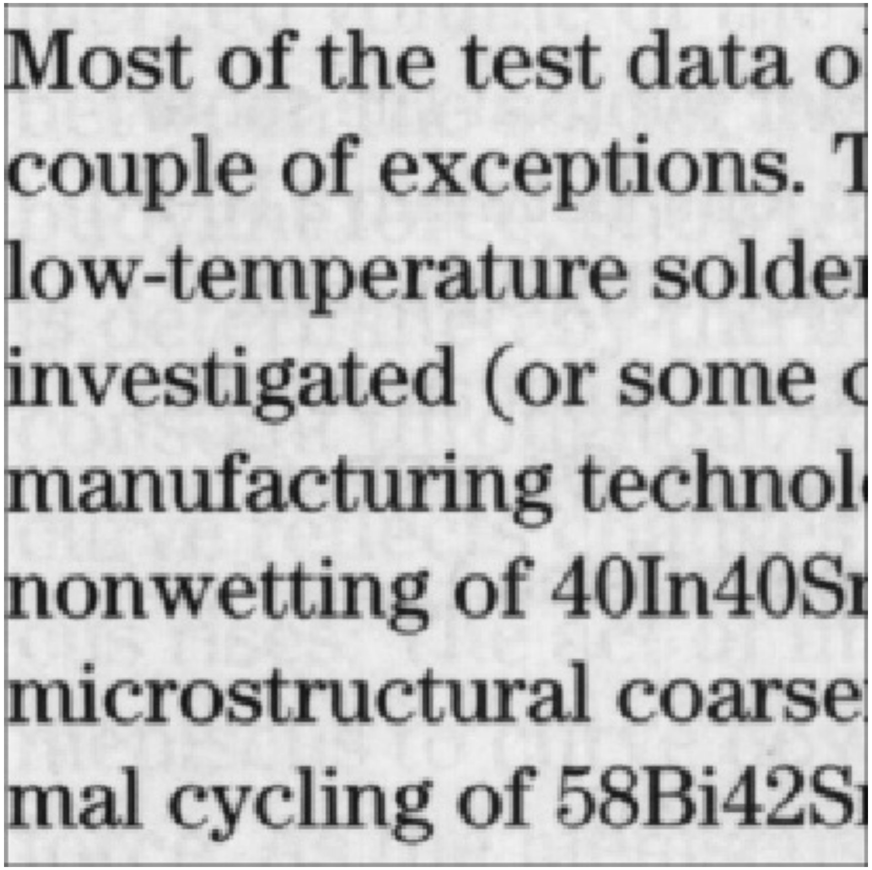


Super-resolution

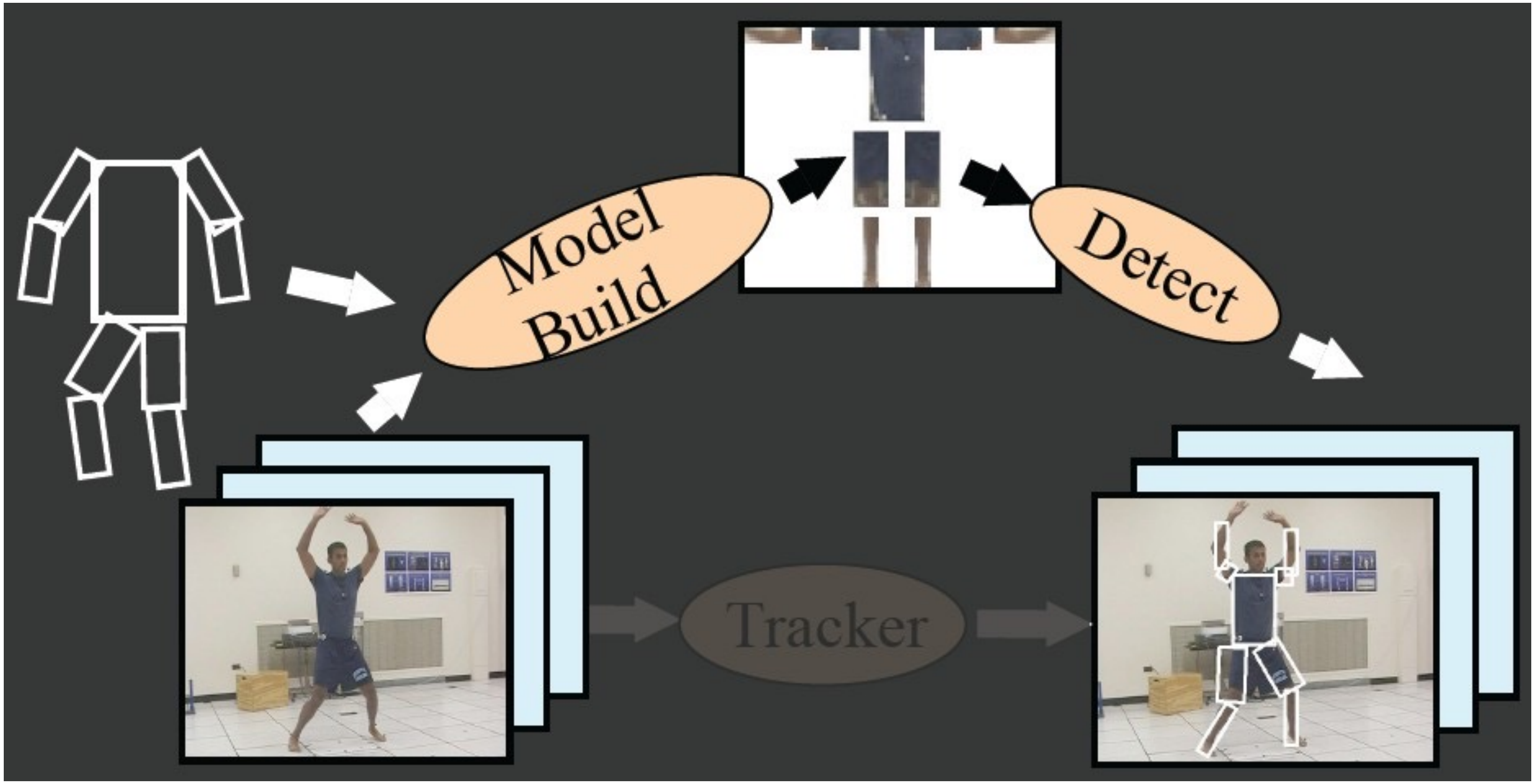
Each of these images looks like this:

Super-resolution

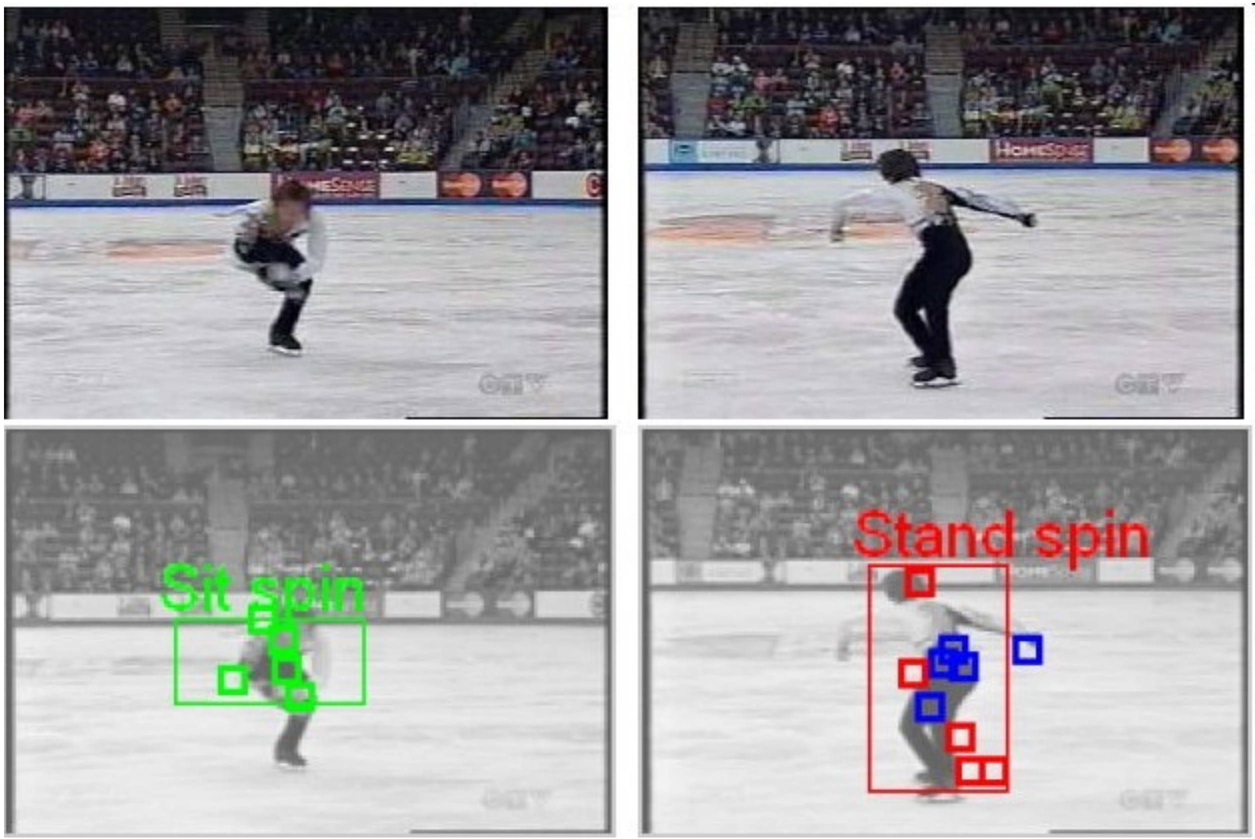
The recovery result:


Most of the test data o
couple of exceptions. T
low-temperature solder
investigated (or some
manufacturing technol
nonwetting of 40In40S
microstructural coarse
mal cycling of 58Bi42S

Recognizing events and activities

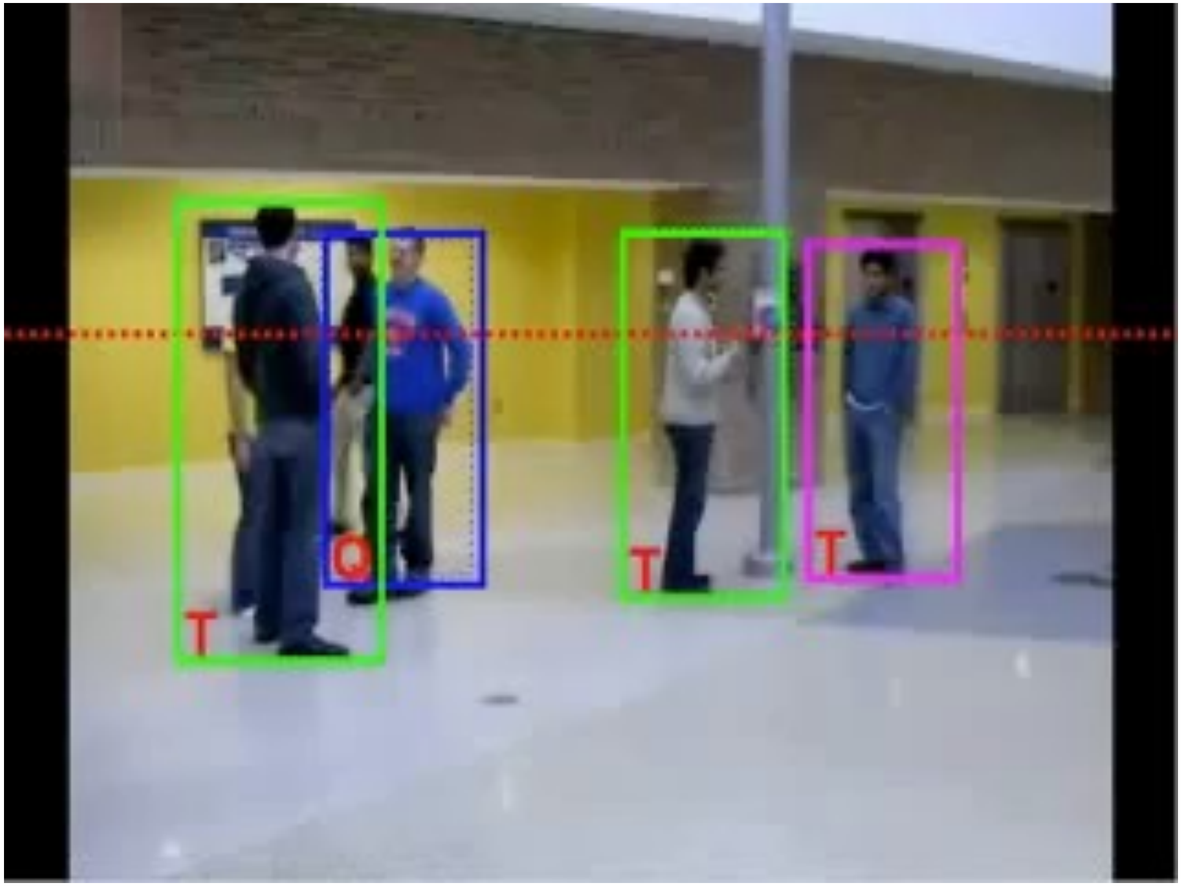


Recognizing events and activities



Recognizing events and activities

Crossing – Talking – Queuing – Dancing – jogging



What we have learned today?

- Optical flow
- Lucas-Kanade method
- Pyramids for large motion
- Horn-Schunk method
- Applications