机器视觉测量与建模

Machine vision based surveying and modelling

李明磊

南京航空航天大学 电子信息工程学院

E-mail: minglei_li@nuaa.edu.cn

课后学习：Feature tracking

- Identify features and track them over video
  - Small difference between frames
  - potential large difference overall

- Standard approach:

  KLT (Kanade-Lukas-Tomasi)

Good features to track

- Use same window in feature selection as for tracking itself

with $\mathbf{M} = \iint_W \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} w(x,y)\mathrm{d}x\mathrm{d}y$

- Compute motion assuming it is small

$$min \iint_W (I + \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \Delta - J)^2 w(x,y)\mathrm{d}x\mathrm{d}y$$

differentiate: $\iint_W 2 \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} (I + \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} \Delta - J) w(x,y)\mathrm{d}x\mathrm{d}y$

$$\iint_W \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} \end{bmatrix} w(x,y)\mathrm{d}x\mathrm{d}y\Delta = \iint_W \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} (J-I) w(x,y)\mathrm{d}x\mathrm{d}y$$

Affine is also possible, but a bit harder (6x6 in stead of 2x2)

Example

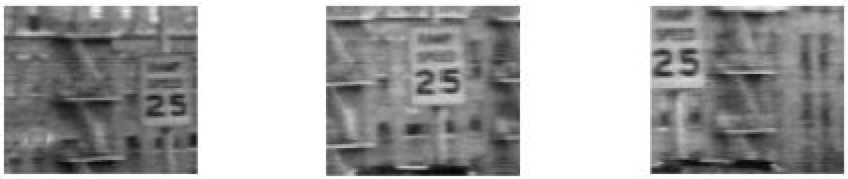
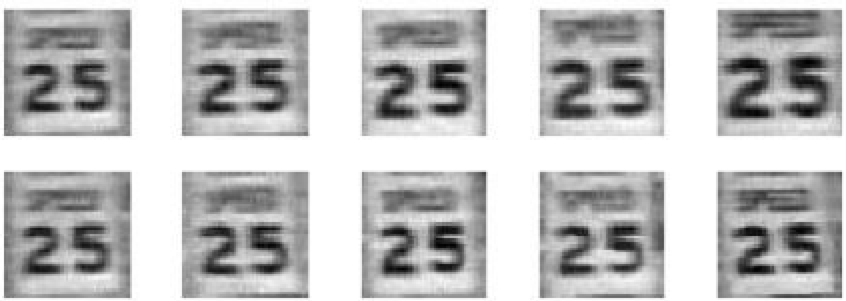
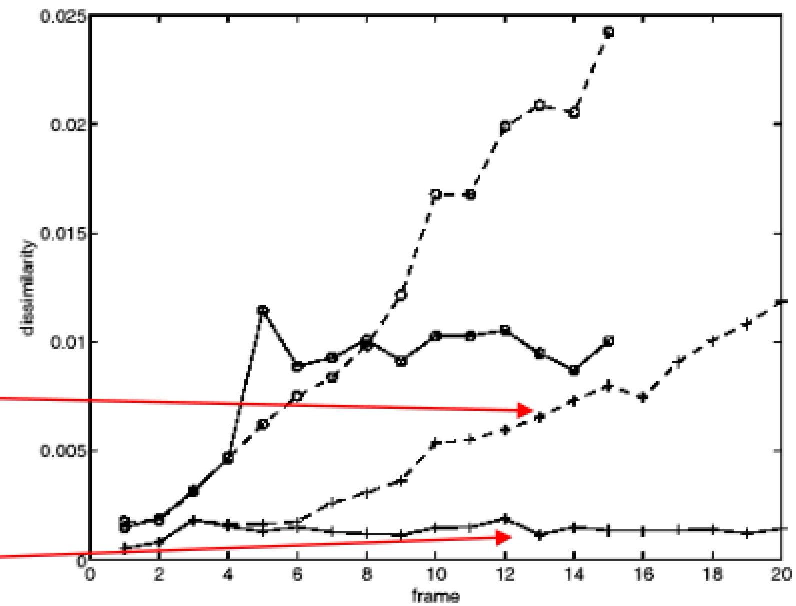


Figure 1: Three frame details from Woody Allen's *Manhattan*. The details are from the 1st, 11th, and 21st frames of a subsequence from the movie.





Simple displacement is sufficient between consecutive frames, but not to compare to reference template

Example

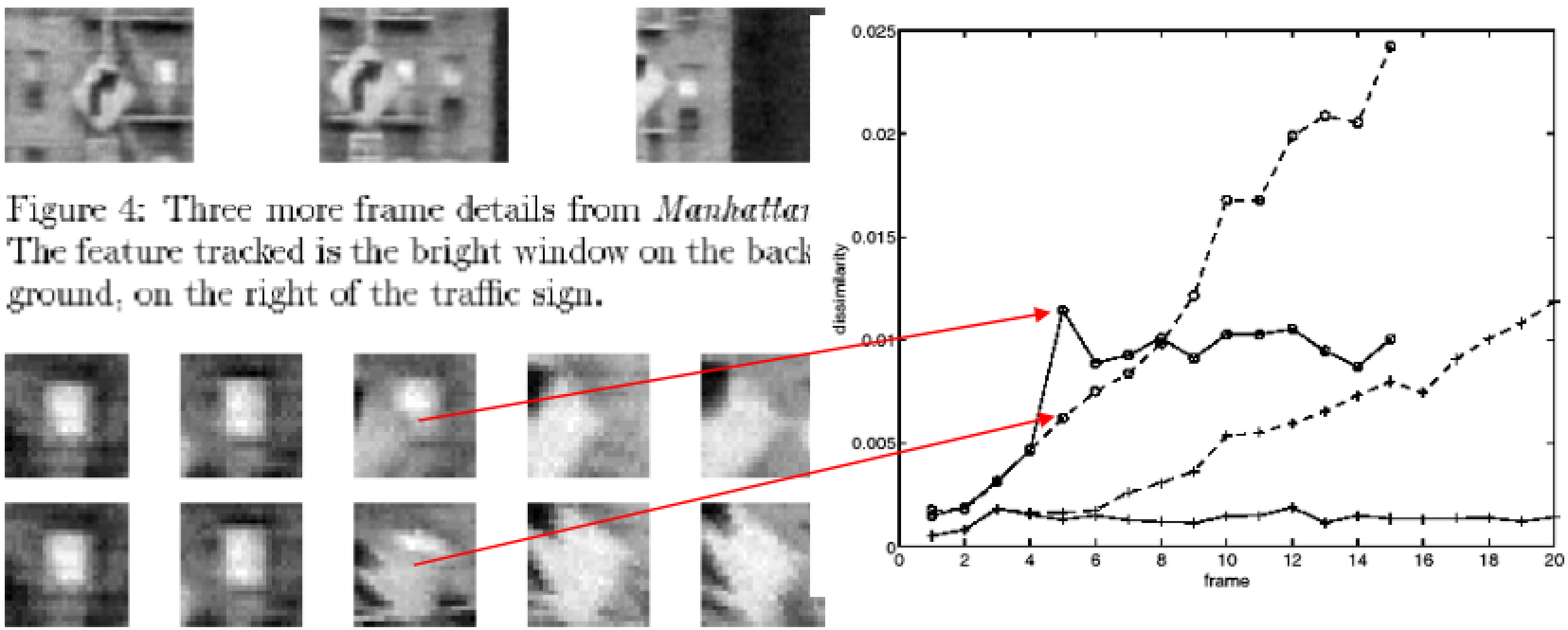
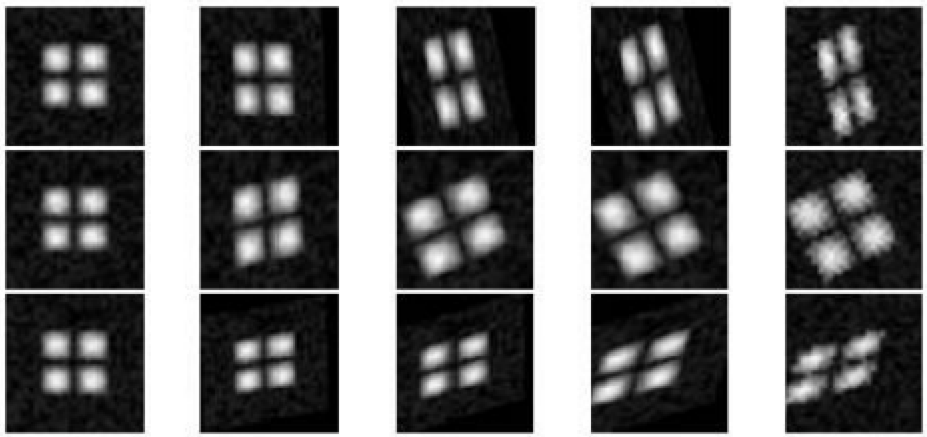
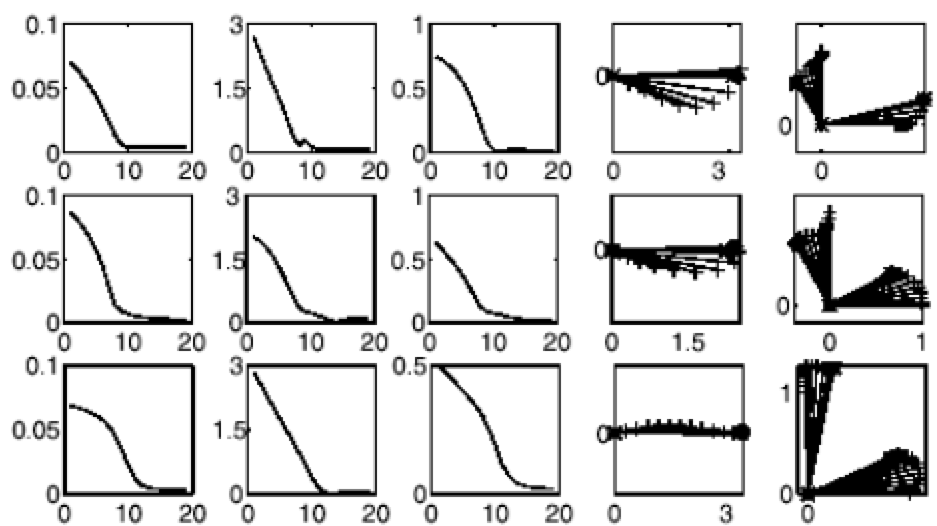


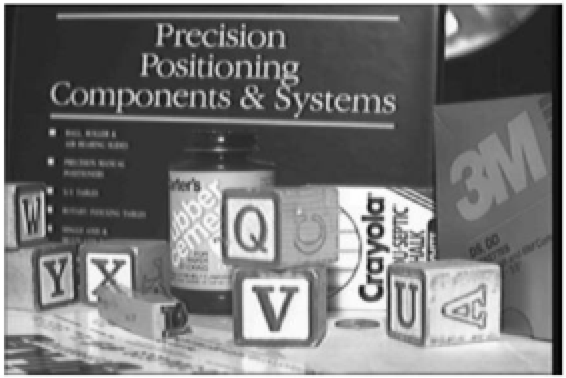
Figure 4: Three more frame details from Manhattan
The feature tracked is the bright window on the back
ground, on the right of the traffic sign.
0.025
0.02
0.015
0.01
0.005
0
dissimilarity
0
2
4
6
8
10
12
14
16
18
20
frame

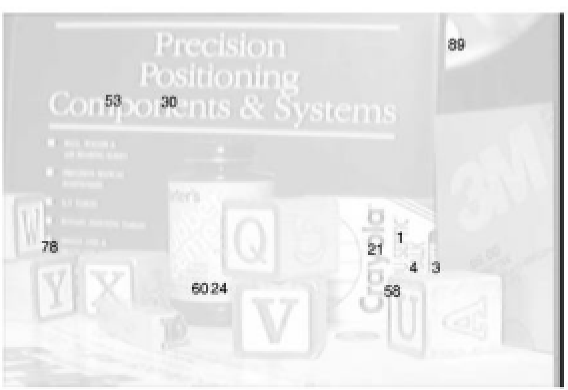
Synthetic example





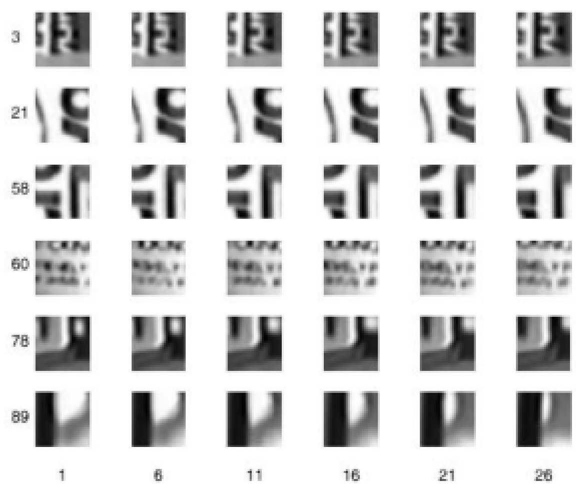
Good features to keep tracking

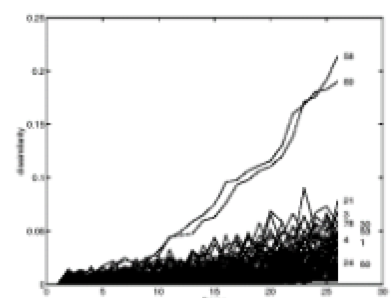


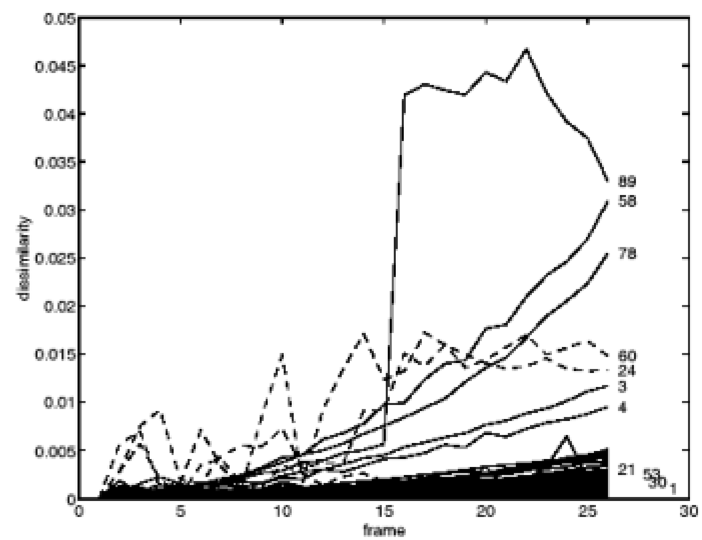


Perform affine alignment between first and last frame

Stop tracking features with too large errors







Optical flow

- Brightness constancy assumption

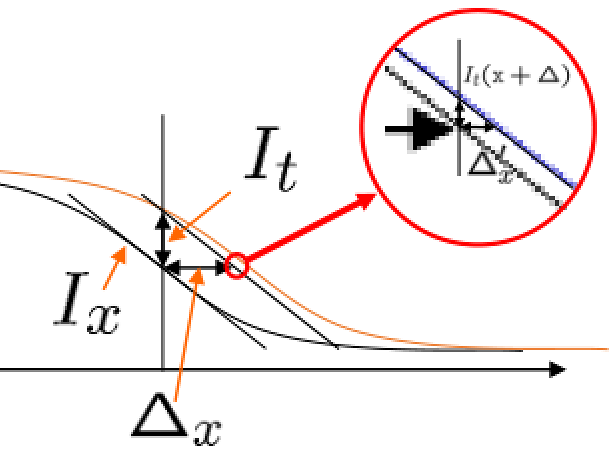
$$I(x+\Delta_x, y+\Delta_y, t+1) = I(x,y,t)$$

$$I(x+u, y+v, t+1) = I(x,y,t) + I_x\Delta_x + I_y\Delta_y + I_t \quad \text{(small motion)}$$

$$I_x\Delta_x + I_y\Delta_y + I_t = 0$$

- 1D example

$$I_x\Delta_x + I_t = 0$$



possibility for iterative refinement

Optical flow

- Brightness constancy assumption

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

$$I(x+u, y+v, t+1) = I(x, y, t) + I_x \Delta_x + I_y \Delta_y + I_t \quad \text{(small motion)}$$

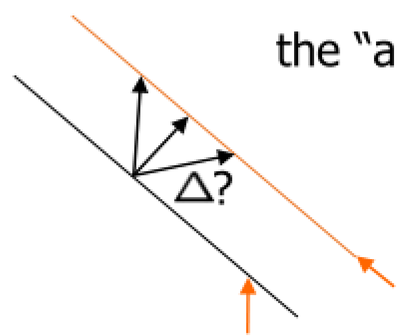
$$I_x \Delta_x + I_y \Delta_y + I_t = 0$$

- 2D example

$$I_x \Delta_x + I_y \Delta_y + I_t = 0$$

(1 constraint)

Δ_x, Δ_y (2 unknowns)



Optical flow

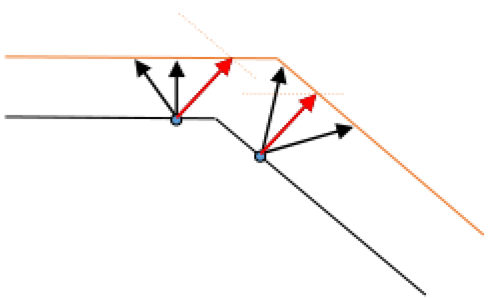
- How to deal with aperture problem?

$$R_x\Delta_x + R_y\Delta_y + R_t = 0 \quad G_x\Delta_x + G_y\Delta_y + G_t = 0 \quad B_x\Delta_x + B_y\Delta_y + B_t = 0$$

(3 constraints if color gradients are different)

Assume neighbors have same displacement

$$I_x(\mathrm{x})\Delta_x + I_y(\mathrm{x})\Delta_y + I_t(\mathrm{x}) = 0 \quad I_x(\mathrm{x}')\Delta_x + I_y(\mathrm{x}')\Delta_y + I_t(\mathrm{x}') = 0 \quad \ldots$$



Lucas-Kanade

Assume neighbors have same displacement

least-squares:

$$\begin{bmatrix} I_x(\mathrm{x}) & I_y(\mathrm{x}) \\ I_x(\mathrm{x}) & I_y(\mathrm{x}) \\ I_x(\mathrm{x}) & I_y(\mathrm{x}) \end{bmatrix} \Delta = \begin{bmatrix} -I_t(\mathrm{x}) \\ -I_t(\mathrm{x}') \\ -I_t(\mathrm{x}'') \end{bmatrix}$$

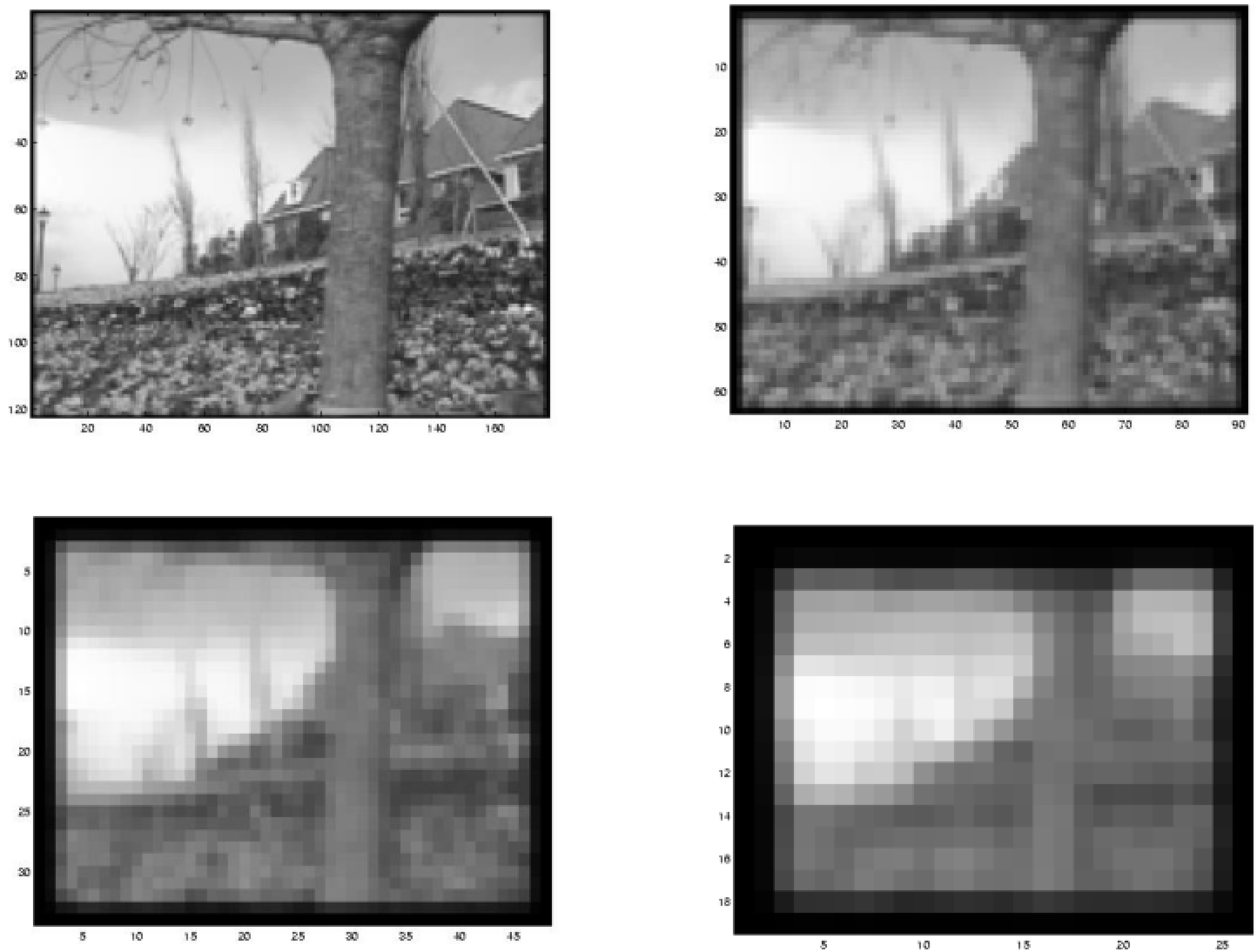
$$\mathbf{A}\Delta = \mathbf{b}$$

Revisiting the small motion assumption

- Is this motion small enough?
  - Probably not—it's much larger than one pixel (2nd order terms dominate)
  - How might we solve this problem?

* From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

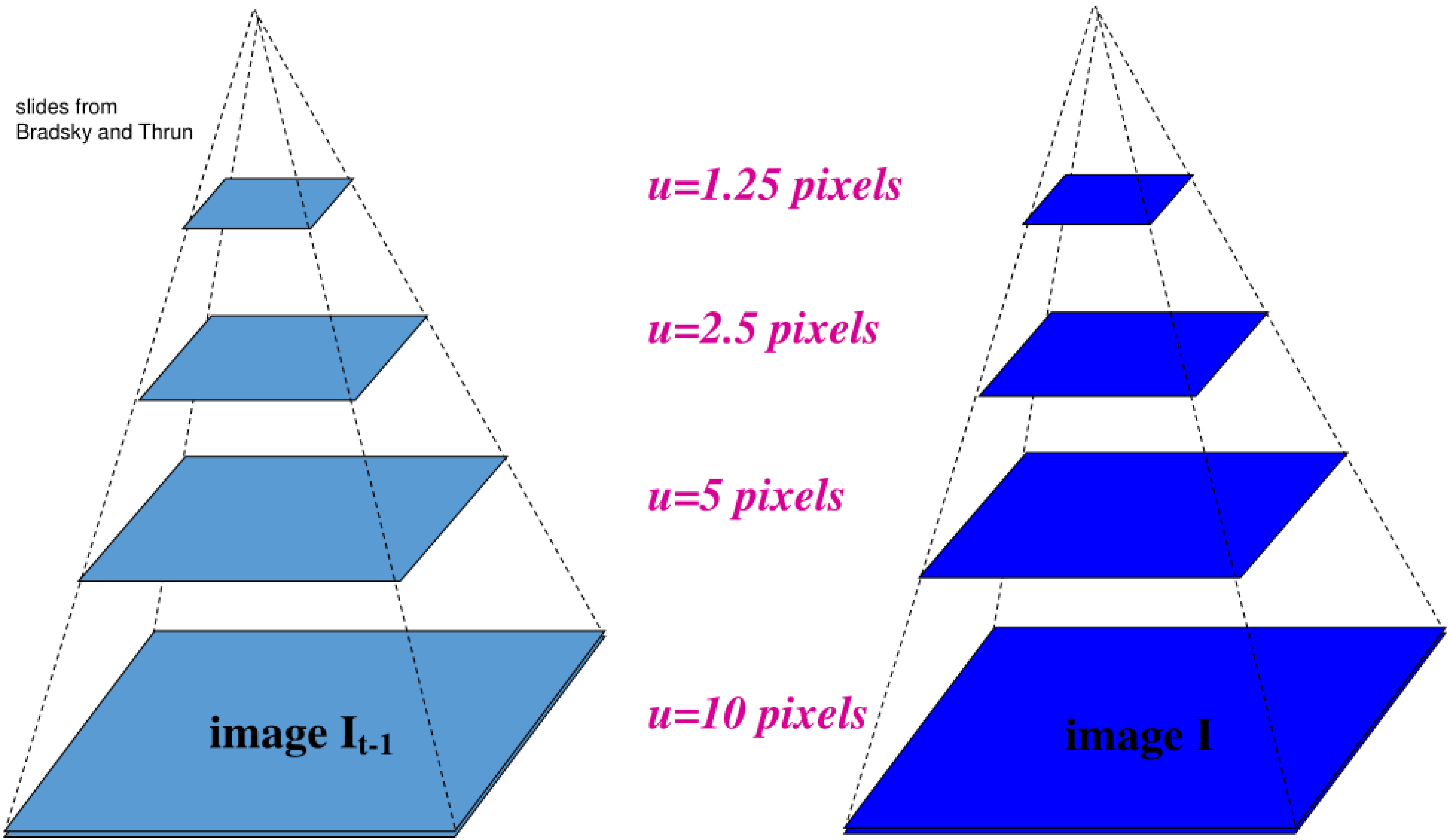
Reduce the resolution!



* From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

Coarse-to-fine optical flow estimation

slides from
Bradsky and Thrun



Gaussian pyramid of image I_{t-1} Gaussian pyramid of image I