# Computer Vision

Spring 2006 15-385,-685

Instructor: S. Narasimhan

Wean 5403

T-R 3:00pm – 4:20pm

Lecture #16

## Announcements

- Homework 4 due today.
- Homework 5 will be out tonight, due in two weeks.
- Use bboard frequently and visit us during OH.

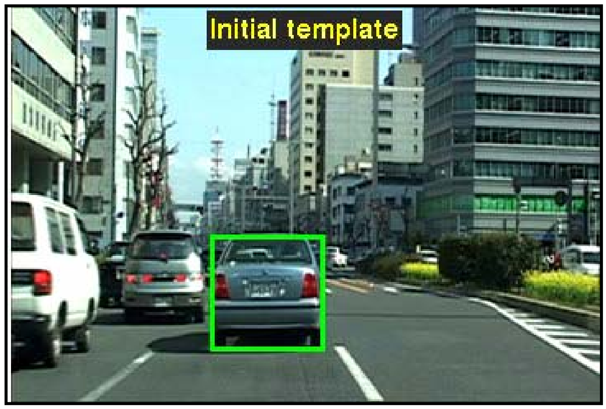
# Optical Flow and Motion

# Lecture #16

## Optical Flow and Motion

- We are interested in finding the movement of scene objects from time-varying images (videos).
- Lots of uses
  - Track object behavior
  - Correct for camera jitter (stabilization)
  - Align images (mosaics)
  - 3D shape reconstruction
  - Special effects

# Tracking – Rigid Objects

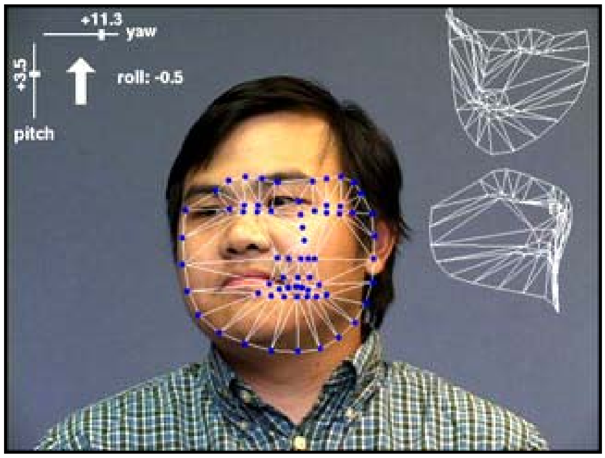



(Simon Baker, CMU)

# Tracking – Non-rigid Objects

(Comaniciu et al, Siemens)

# Face Tracking



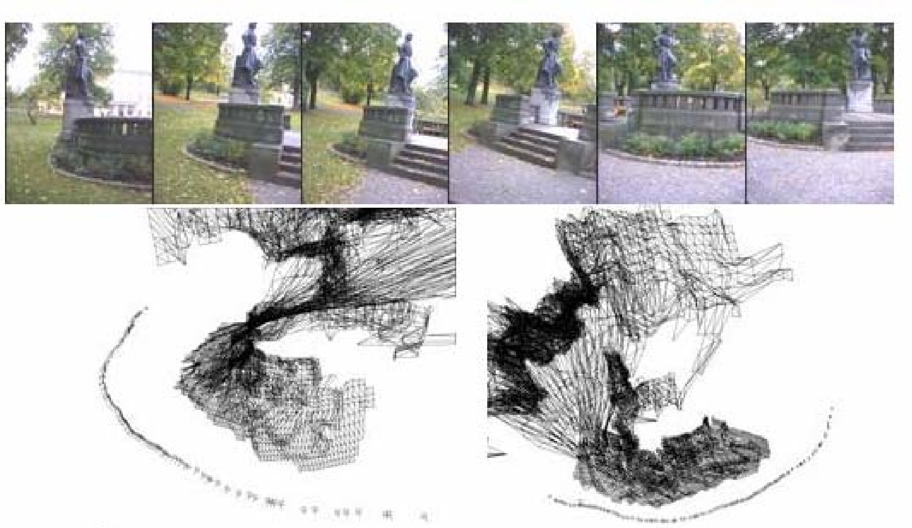

(Simon Baker et al, CMU)

# Applications of Face Tracking

- User Interfaces:
  - Mouse Replacement: Head Pose and Gaze Estimation
  - Automotive: Windshield Displays, Smart Airbags, Driver Monitoring
- Face Recognition:
  - Pose Normalization
  - Model-Based Face Recognition
- Lipreading/Audio-Visual Speech Recognition
- Expression Recognition and Deception Detection
- Rendering and Animation:
  - Expression Animations and Transfer
  - Low-Bandwidth Video Conferencing
  - Audio-Visual Speech Synthesis

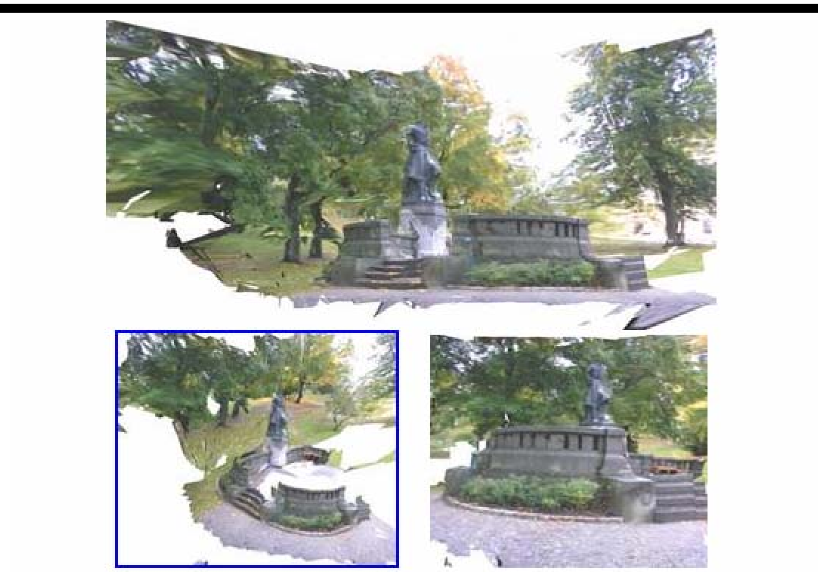
(Simon Baker, CMU)

# 3D Structure from Motion



(David Nister, Kentucky)

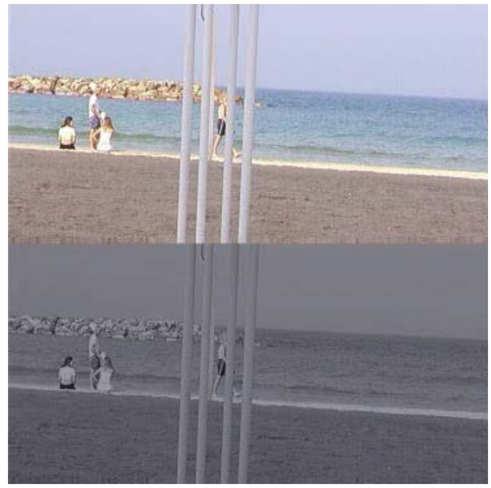
# 3D Structure from Motion



(David Nister, Kentucky)

# Behavior Analysis

Query

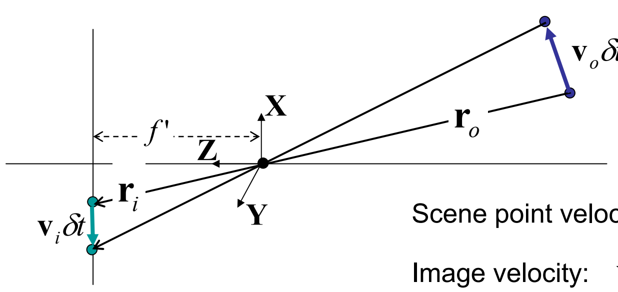


Result

(Michal Irani, Weizmann)

# Motion Field

- Image velocity of a point moving in the scene




Scene point velocity: $\mathbf{v}_o = \dfrac{d\mathbf{r}_o}{dt}$

Image velocity: $\mathbf{v}_i = \dfrac{d\mathbf{r}_i}{dt}$

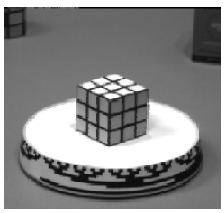
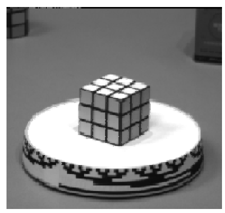
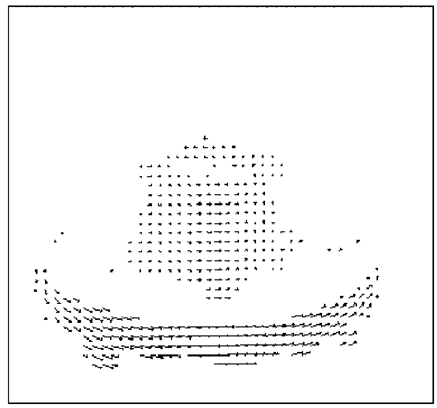
Perspective projection: $\dfrac{1}{f'}\mathbf{r}_i = \dfrac{\mathbf{r}_o}{\mathbf{r}_o \cdot \mathbf{Z}}$

Motion field

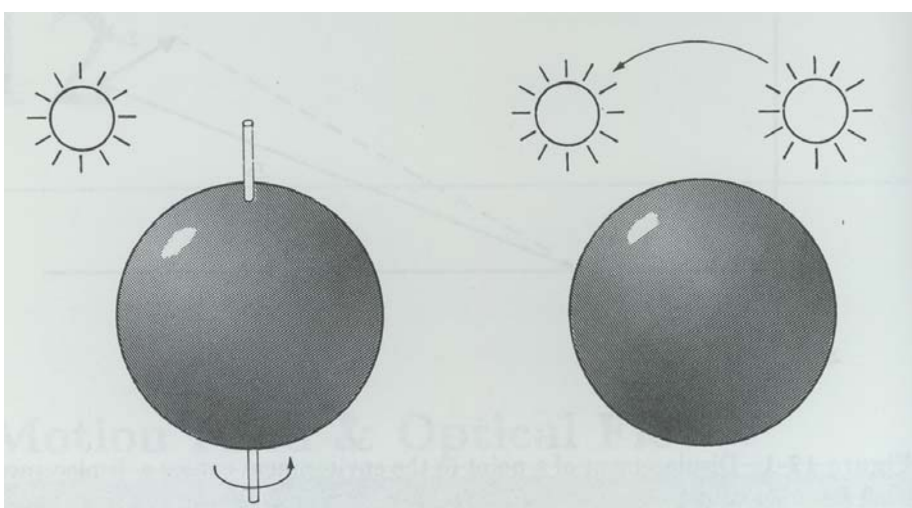
$$\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt} = f'\frac{(\mathbf{r}_o \cdot \mathbf{Z})\mathbf{v}_o - (\mathbf{v}_o \cdot \mathbf{Z})\mathbf{r}_o}{(\mathbf{r}_o \cdot \mathbf{Z})^2} = f'\frac{(\mathbf{r}_o \times \mathbf{v}_o) \times \mathbf{Z}}{(\mathbf{r}_o \cdot \mathbf{Z})^2}$$

# Optical Flow

- Motion of brightness pattern in the image
- **Ideally** Optical flow = Motion field

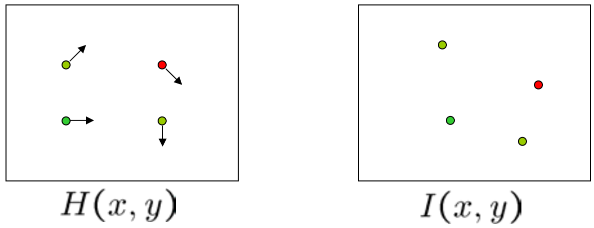
# Optical Flow $\neq$ Motion Field



Motion field exists but no optical flow
(a)

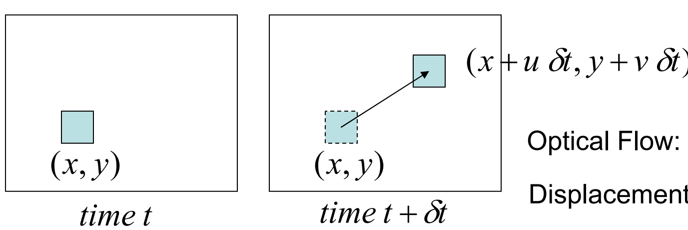
No motion field but shading changes
(b)

# Problem Definition: Optical Flow




- How to estimate pixel motion from image H to image I?
  - Find pixel correspondences
    - Given a pixel in H, look for nearby pixels of the same color in I
- Key assumptions
  - **color constancy**: a point in H looks "the same" in image I
    - For grayscale images, this is **brightness constancy**
  - **small motion**: points do not move very far

# Optical Flow Constraint Equation




Optical Flow: Velocities $(u,v)$

Displacement:

$$(\delta x, \delta y) = (u\ \delta t, v\ \delta t)$$

- Assume brightness of patch remains same in both images:

$$E(x+u\ \delta t, y+v\ \delta t, t+\delta t) = E(x,y,t)$$

- Assume small motion: (Taylor expansion of LHS upto first order)

$$E(x,y,t) + \delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} = E(x,y,t)$$

# Optical Flow Constraint Equation

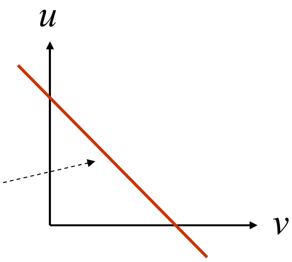
$$\delta x \frac{\partial E}{\partial x} + \delta y \frac{\partial E}{\partial y} + \delta t \frac{\partial E}{\partial t} = 0$$

Divide by $\delta t$ and take the limit $\delta t \to 0$

$$\frac{dx}{dt}\frac{\partial E}{\partial x} + \frac{dy}{dt}\frac{\partial E}{\partial y} + \frac{\partial E}{\partial t} = 0$$

Constraint Equation

$$\boxed{E_x\, u + E_y\, v + E_t = 0}$$

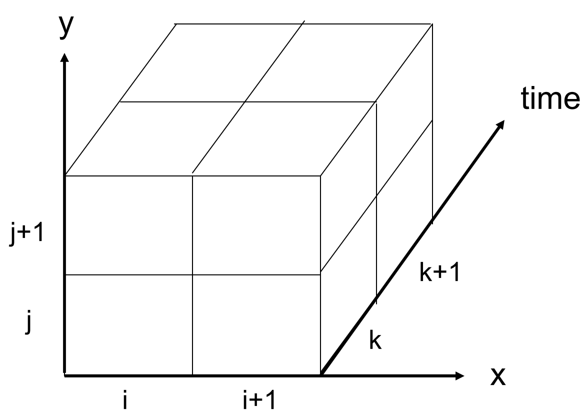



NOTE: $(u,v)$ must lie on a straight line

We can compute $E_x$ , $E_y$ , $E_t$ using gradient operators!

But, (u,v) cannot be found uniquely with this constraint!

# Finding Gradients in X-Y-T



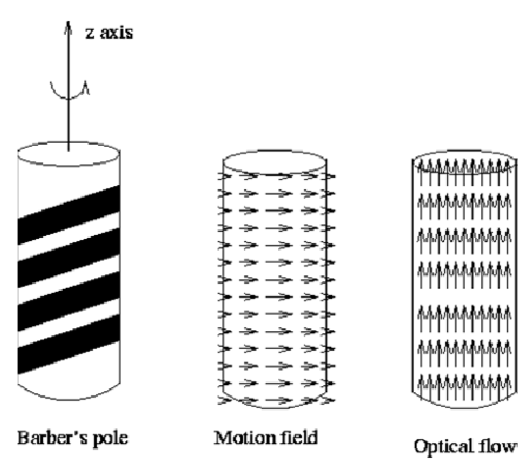

$$E_x = \frac{1}{4\,\delta x}[(E_{i+1,j,k} + E_{i+1,j,k+1} + E_{i+1,j+1,k} + E_{i+1,j+1,k+1})$$
$$-(E_{i,j,k} + E_{i,j,k+1} + E_{i,j+1,k} + E_{i,j+1,k+1})]$$

# Optical Flow Constraint

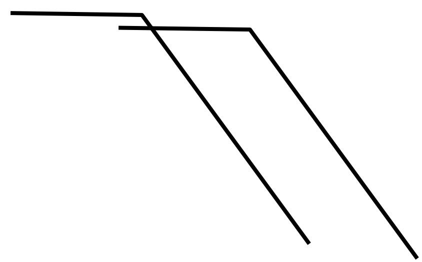
- Intuitively, what does this constraint mean?
  - The component of the flow in the gradient direction is determined
  - The component of the flow parallel to an edge is unknown

# Optical Flow Constraint

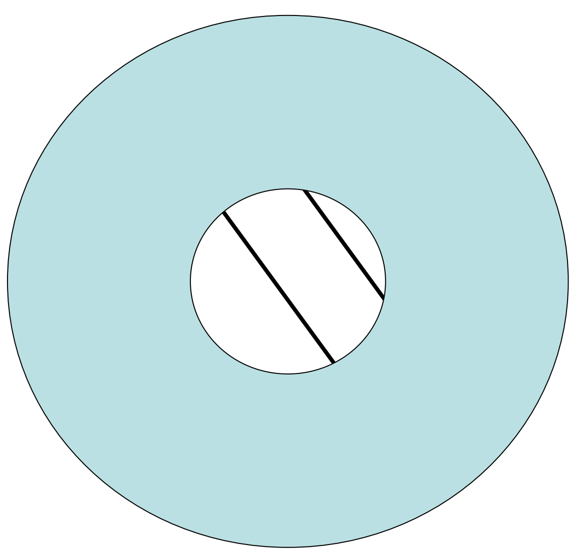
Barber pole illusion



## Aperture Problem



## Aperture Problem



# Computing Optical Flow

- Formulate Error in Optical Flow Constraint:

$$e_c = \iint_{image} (E_x u + E_y v + E_t)^2 \, dx \, dy$$

- We need additional constraints!
- Smoothness Constraint (as in shape from shading and stereo):

Usually motion field varies smoothly in the image.
So, penalize departure from smoothness:

$$e_s = \iint_{image} (u_x^2 + u_y^2) + (v_x^2 + v_y^2) \;\; dx \, dy$$

- Find (u,v) at each image point that MINIMIZES:

$$e = e_s + \lambda e_c$$

weighting factor

# Discrete Optical Flow Algorithm

Consider image pixel $(i, j)$

- Departure from Smoothness Constraint:

$$s_{ij} = \frac{1}{4}[(u_{i+1,j} - u_{i,j})^2 + (u_{i,j+1} - u_{i,j})^2 + (v_{i+1,j} - v_{i,j})^2 + (v_{i,j+1} - v_{i,j})^2]$$

- Error in Optical Flow constraint equation:

$$c_{ij} = (E^{ij}{}_x \, u_{ij} + E^{ij}{}_y \, v_{ij} + E^{ij}{}_t)^2$$

- We seek the set $\{u_{ij}\}$ & $\{v_{ij}\}$ that minimize:

$$e = \sum_i \sum_j (s_{ij} + \lambda \, c_{ij})$$

NOTE: $\{u_{ij}\}$ & $\{v_{ij}\}$ show up in more than one term

# Discrete Optical Flow Algorithm

- Differentiating $e$ w.r.t $v_{kl}$ & $u_{kl}$ and setting to zero:

$$\frac{\partial e}{\partial u_{kl}} = 2\,(u_{kl} - \overline{u_{kl}}) + 2\lambda\,(E_x^{kl}u_{kl} + E_y^{kl}v_{kl} + E_t^{kl})E_x^{kl} = 0$$

$$\frac{\partial e}{\partial v_{kl}} = 2\,(v_{kl} - \overline{v_{kl}}) + 2\lambda\,(E_x^{kl}u_{kl} + E_y^{kl}v_{kl} + E_t^{kl})E_y^{kl} = 0$$

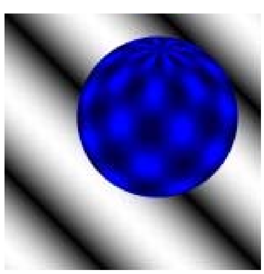
- $\overline{v_{kl}}$ & $\overline{u_{kl}}$ are averages of $(u,v)$ around pixel $(k,l)$

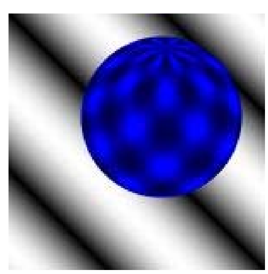
Update Rule:

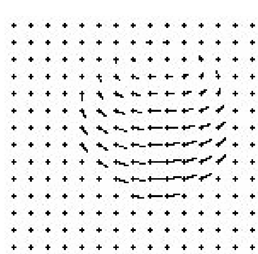
$$u_{kl}^{n+1} = \overline{u_{kl}^{n}} - \frac{E_x^{kl}\,\overline{u_{kl}^{n}} + E_y^{kl}\,\overline{v_{kl}^{n}} + E_t^{kl}}{1+\lambda\,[(E_x^{kl})^2 + (E_y^{kl})^2]}E_x^{kl}$$

$$v_{kl}^{n+1} = \overline{v_{kl}^{n}} - \frac{E_x^{kl}\,\overline{u_{kl}^{n}} + E_y^{kl}\,\overline{v_{kl}^{n}} + E_t^{kl}}{1+\lambda\,[(E_x^{kl})^2 + (E_y^{kl})^2]}E_y^{kl}$$

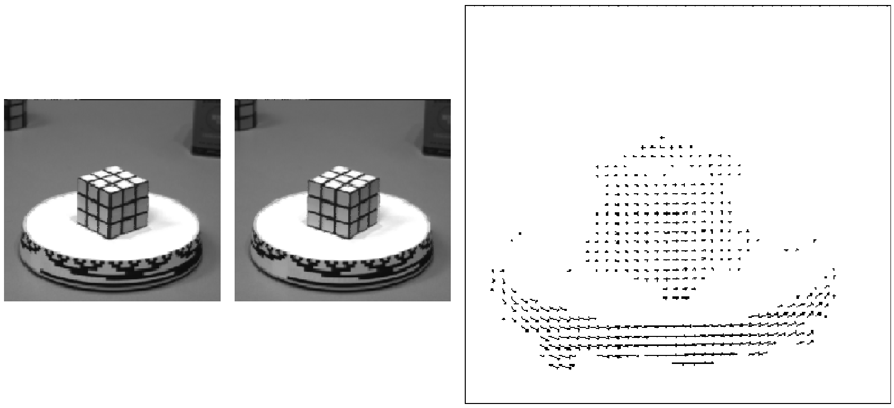
# Example



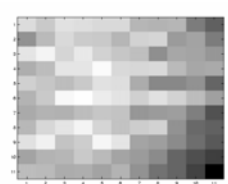




# Optical Flow Result

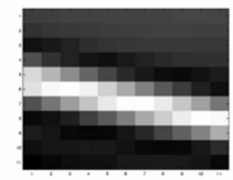


# Low Texture Region - Bad

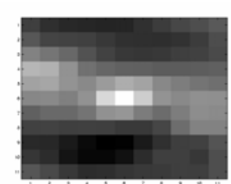
– gradients have small magnitude

# Edges – so,so (aperture problem)



– large gradients, all the same

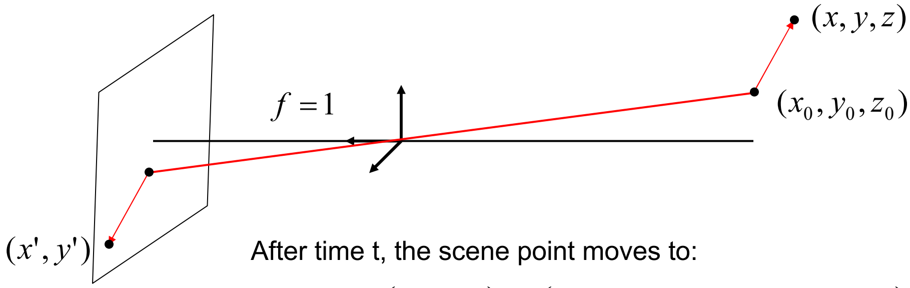
# High Textured Region - Good



– gradients are different, large magnitudes

# Focus of Expansion (FOE)

- Motion of object = - (Motion of Sensor)
- For a given translatory motion and gaze direction, the world seems to flow out of one point (FOE).

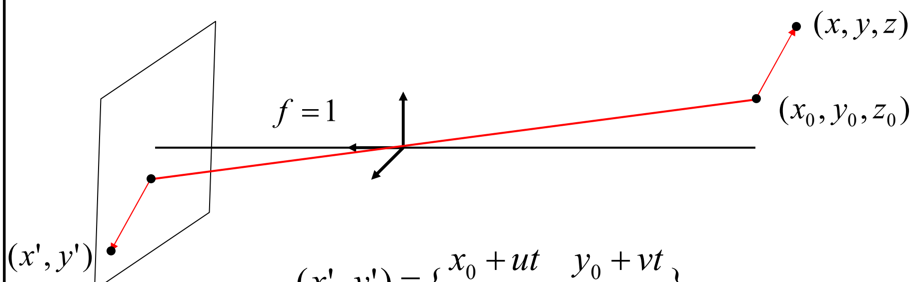



After time t, the scene point moves to:

$$(x, y, z) \equiv (x_0 + ut, y_0 + vt, z_0 + wt)$$

$$(x', y') \equiv \{\frac{x_0 + ut}{z_0 + wt}, \frac{y_0 + vt}{z_0 + wt}\}$$

# Focus of Expansion (FOE)




$$(x', y') \equiv \{\frac{x_0 + ut}{z_0 + wt}, \frac{y_0 + vt}{z_0 + wt}\}$$

- As t varies the image point moves along a straight line in the image
- Focus of Expansion: Lets backtrack time or $(t \rightarrow -\infty)$

$$\boxed{(x', y') \equiv \{\frac{u}{w}, \frac{v}{w}\}}$$

# Focus of Expansion (FOE) - Example

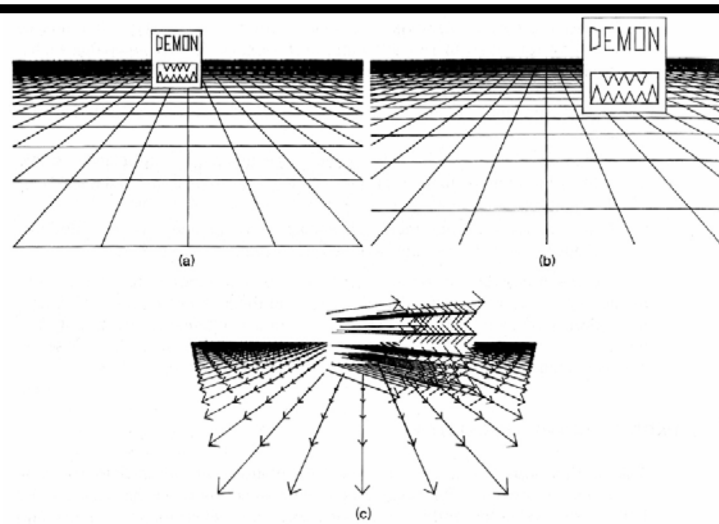


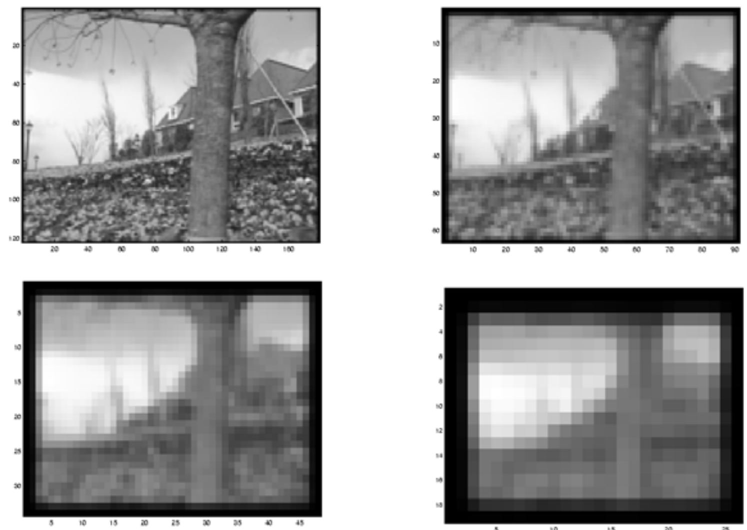

Fig. 7.3 FOE for rectilinear observer motion. (a) An image. (b) Later image. (c) Flow shows different FOEs for static floor and moving object.

http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/LIB/bandb7_12.pdf

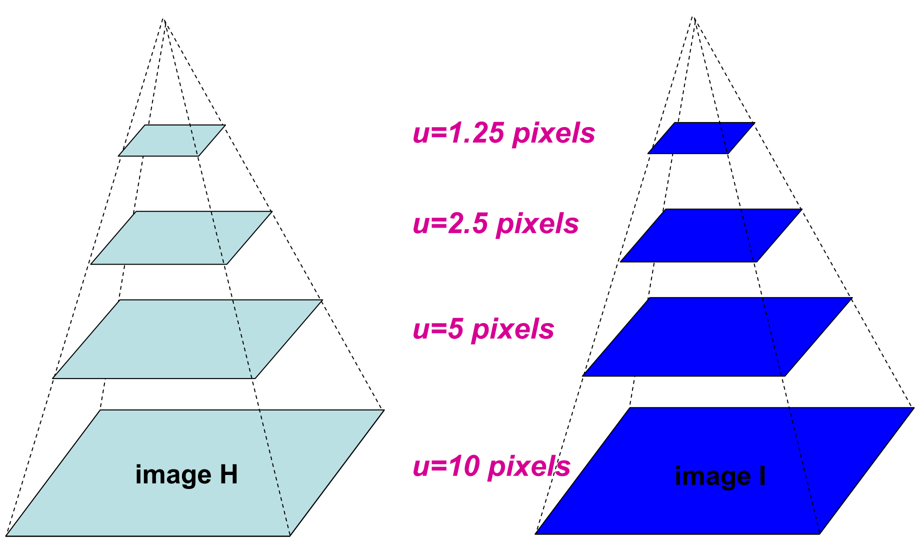
# Revisiting the Small Motion Assumption

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

# Reduce the Resolution!

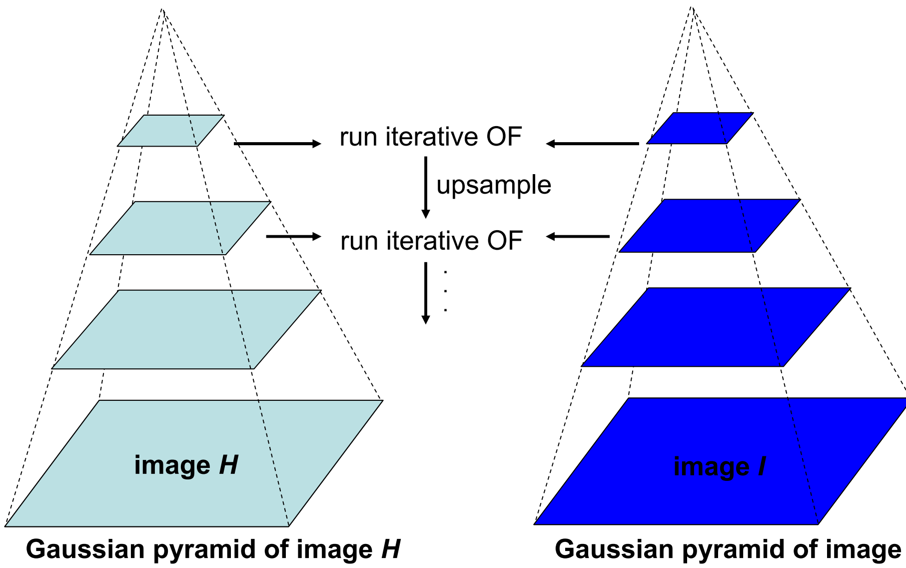


# Coarse-to-fine Optical Flow Estimation

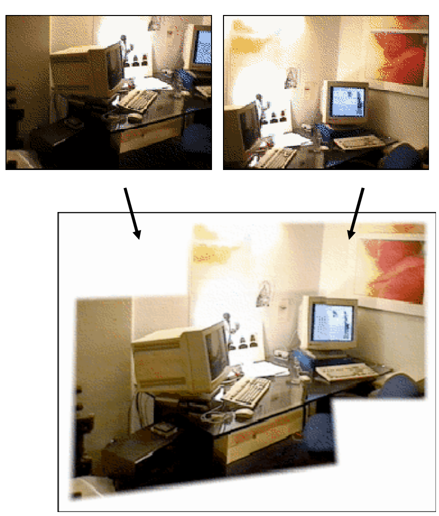


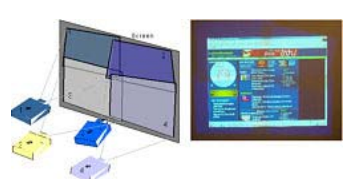

**Gaussian pyramid of image *H***    **Gaussian pyramid of image *I***

# Coarse-to-fine Optical Flow Estimation



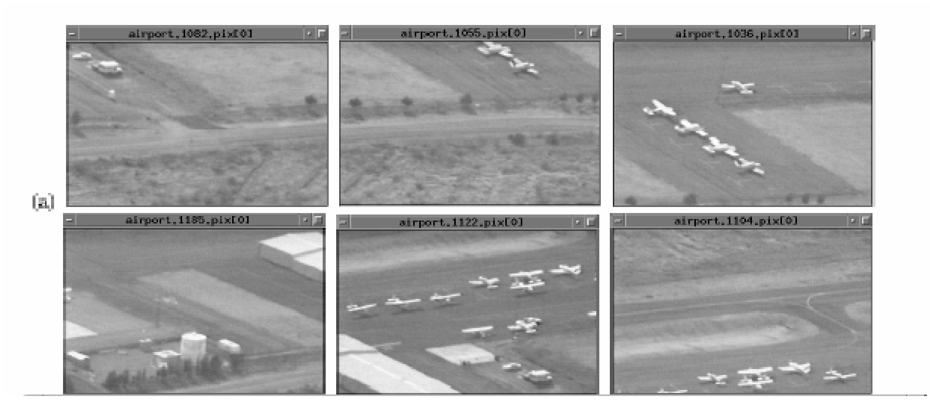

# Image Alignment





- Goal: Estimate single *(u,v)* translation (transformation) for entire image

# Mosaicing




(Michal Irani, Weizmann)

# Mosaicing


1. Static background mosaic of an airport video clip.
(a) A few representative frames from the minute-long video clip. The video shows an airport being imaged from the air with a moving camera. The scene itself is static (i.e., no moving objects). (b) The static background mosaic image which provides an extended view of the entire scene imaged by the camera in the one-minute video clip.

(Michal Irani, Weizmann)

# Next Class

- Structured Light and Range Imaging
- Reading → Notes