Computer Vision, Lectures 15,16

Professor Hager
http://www.cs.jhu.edu/~hager

Outline for Today

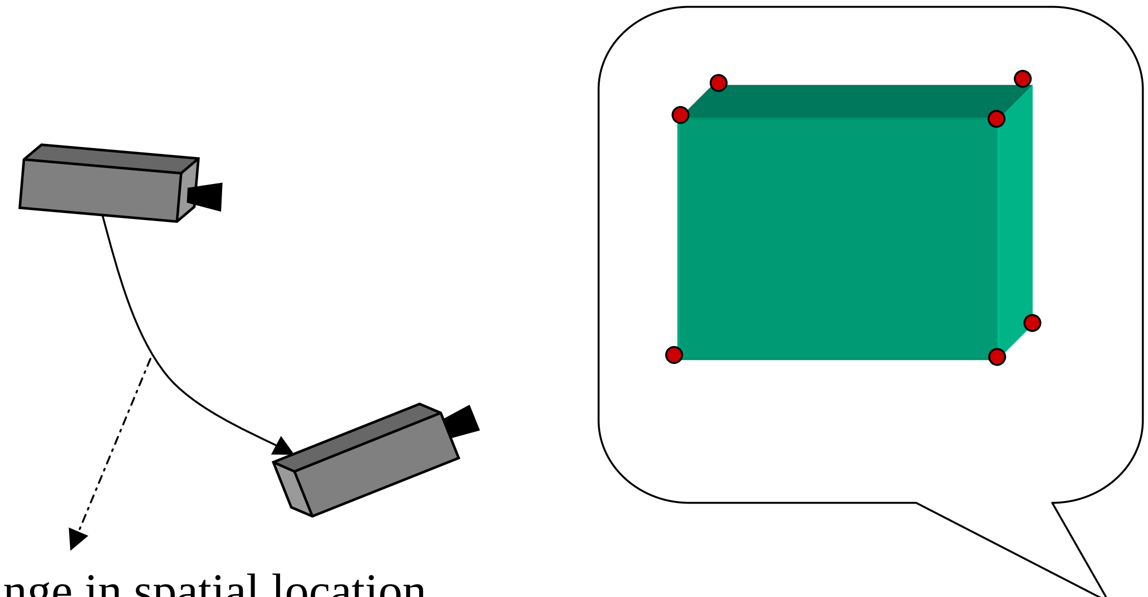
From Stereo to Motion

The motion field and optical flow

Your next assignment

DYNAMIC VISION: THE PROBLEM

MOVING CAMERAS ARE LIKE STEREO



The change in spatial location
between the two cameras (the “motion”)

Locations of
points on the object
(the “structure”)

Result of Tomasi's SFM algorithm

Simplify camera model (linear)
Use large amounts of data (multiple views, many points)
Solve correspondence using tracking methods

SOME CORE PROBLEMS

- Image (2D) motion estimation
  - motion detection
  - optical flow calculation
- 3D motion estimation
  - ego (self) motion
  - target motion
  - structure from motion differential
  - matching methods
- Structured motion recovery
  - visual tracking

W/O loss of generality, we will assume a unit focal length, metric camera

What is Visual Motion?

$^{c}p = {}^{w}p - Tr$

1. A body moving in space experiences
   - translation (Tr)
   - rotation (R)

2. We are interested in time change
   - trans. velocity (T = dTr/dt)
   - rot. velocity (ω = "dR/dt")

3. Other rigid bodies appear to move in the opposite direction

What is Visual Motion?

1. A body moving in space experiences
   - translation (Tr)
   - rotation (R)

2. We are interested in time change
   - trans. velocity ($T = dTr/dt$)
   - rot. velocity (ω = "dR/dt")

3. Other rigid bodies appear to move in the opposite direction

4. Points on those bodies project into camera images

$${}^{c}p = {}^{w}p - Tr$$

$$u = x/z$$
$$v = y/z$$

What is Visual Motion?

1. A body moving in space experiences
   - translation (Tr)
   - rotation (R)

2. We are interested in time change
   - trans. velocity (T = dTr/dt)
   - rot. velocity (ω = “dR/dt”)

3. Other rigid bodies appear to move in the opposite direction

4. Points on those bodies project into camera images

$$^{c}x = {}^{w}x - Tr_x \rightarrow dx/dt = -T_x$$
$$^{c}y = {}^{w}y - Tr_y \rightarrow dy/dt = -T_y$$
$$^{c}z = {}^{w}z - Tr_z \rightarrow dz/dt = -T_z$$

$$du/dt = (dx/dt\ {}^{c}z - dz/dt\ {}^{c}x)/{}^{c}z^2$$

$$u = x/z$$
$$v = y/z$$

What is Visual Motion?

1. A body moving in space experiences
   - translation (Tr)
   - rotation (R)

2. We are interested in time change
   - trans. velocity (T = dTr/dt)
   - rot. velocity (ω = "dR/dt")

3. Other rigid bodies appear to move in the opposite direction

4. Points on those bodies project into camera images

$${}^{c}x = {}^{w}x - Tr_x \rightarrow dx/dt = -T_x$$
$${}^{c}y = {}^{w}y - Tr_y \rightarrow dy/dt = -T_y$$
$${}^{c}z = {}^{w}z - Tr_z \rightarrow dz/dt = -T_z$$

$$du/dt = (dx/dt\ {}^{c}z - dz/dt\ {}^{c}x)/{}^{c}z^2$$
$$du/dt = -(T_x + u\ T_z)/z$$

$$u = x/z$$
$$v = y/z$$

What is Visual Motion?

$^{c}x = {}^{w}x - Tr_x \rightarrow dx/dt = -T_x$

$^{c}y = {}^{w}y - Tr_y \rightarrow dy/dt = -T_y$

$^{c}z = {}^{w}z - Tr_z \rightarrow dz/dt = -T_z$

$du/dt = (dx/dt\ {}^{c}z - dz/dt\ {}^{c}x)/{}^{c}z^2$

$du/dt = -(T_x + u\ T_z)/z$

$dv/dt = -(T_y + v\ T_z)/z$

1. A body moving in space experiences
   - translation (Tr)
   - rotation (R)

2. We are interested in time change
   - trans. velocity (T = dTr/dt)
   - rot. velocity (ω = “dR/dt”)

3. Other rigid bodies appear to move in the opposite direction

4. Points on those bodies project into camera images

$u = x/z$

$v = y/z$

What is Visual Motion?

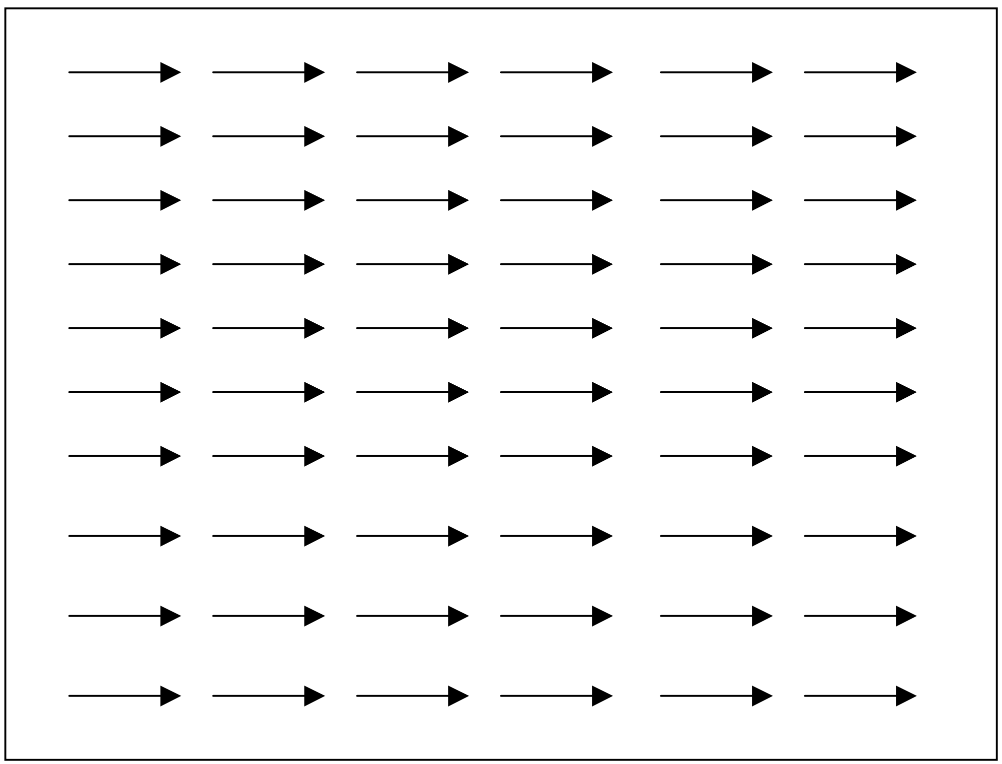
$$du/dt = -(T_x + u\,T_z)/z$$
$$dv/dt = -(T_y + v\,T_z)/z$$

This expression defines the motion field for the image under pure translation

Note that, given a velocity vector, we can determine precisely what the motion field is *up to a scale factor related to the depths of the points that are observed*

THE MOTION FIELD

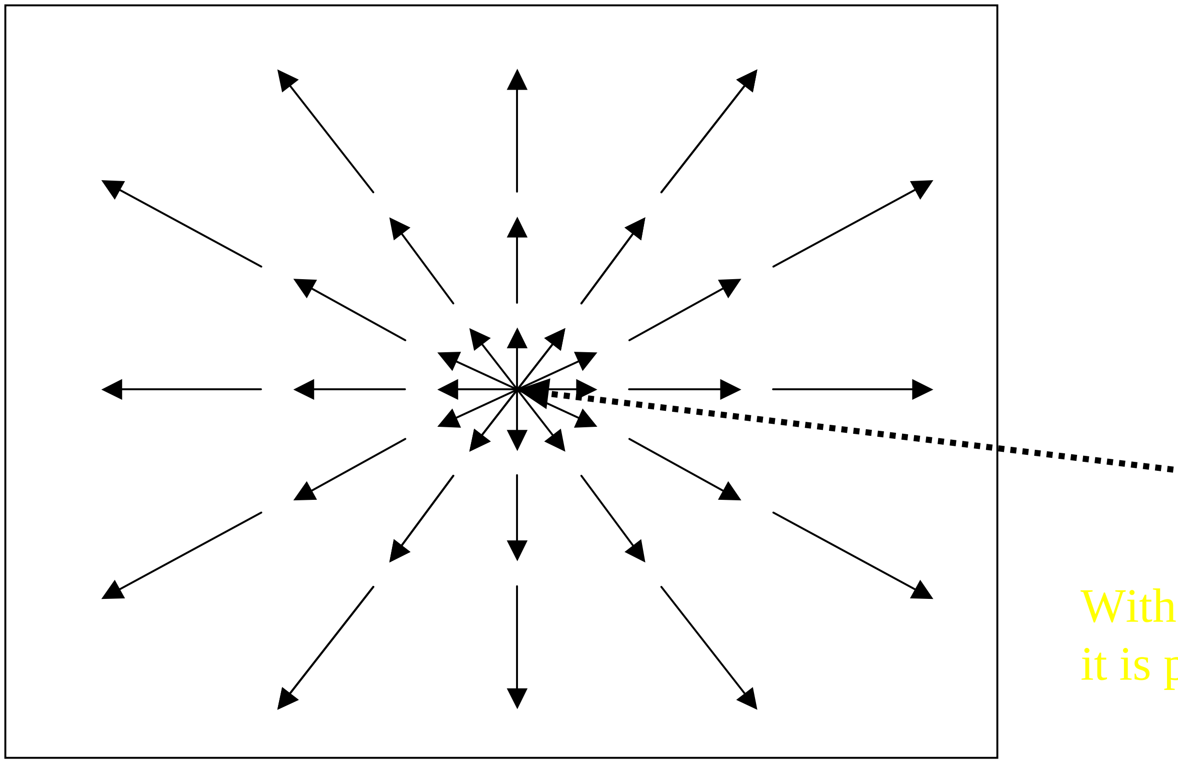
The “instantaneous” velocity of points in an image



TRANSLATION
(OR ROTATION)
TO THE LEFT

THE MOTION FIELD

The "instantaneous" velocity of points in an image



LOOMING

The *focus of expansion*

With just this information it is possible to calculate:

1. Direction of motion
2. Time to collision

Calculations

$du/dt = -(T_x + u\,T_z)/z$
$dv/dt = -(T_y + v\,T_z)/z$

- Define $u_0 = T_x/T_z$ ($T_z \neg= 0$); $v_0 = T_y/T_z$
  - note $(u_0, v_0, 1)$ is direction of translation
- $u' = (u – u_0)\,T_z/Z$; $v' = (v – v_0)\,T_z/Z \rightarrow p' = (p – p_0)\,T_z/Z$
  - therefore, the field is radial and rooted at $(T_x/T_z, T_y/T_z) = p_0$
- Z/T_z is the number of seconds until the point crosses the image plane (“time to collision”) *for this point*
  - $Z/T_z = (u-u_0)/u'$; $Z/T_z = (v – v_0)/v'$

Rotational Motion

- First, consider a 2D rotation $R(\theta) = [\cos(\theta), -\sin(\theta); \sin(\theta), \cos(\theta)]$;
  - ${}^{c}p = R(\theta)\ {}^{w}p$
  - ${}^{c}\ dp/dt = d\ R(\theta)\ dt\ {}^{w}p + R(\theta)\ d\ {}^{w}p/dt$
- If θ is a time-varying function, then
  $dR/d\theta = \omega\ [-\sin(\theta), -\cos(\theta); \cos(\theta), -\sin(\theta)] = [0, -\omega; \omega, 0]\ R(\theta)$
  - Note that the differential rotation creates the instantaneous motion vector of the point
  - We only care about case $\theta = 0$
- In 3D it's essentially the same story
  - $dp/dt = S\ p$
  - $S = [0, -\omega_z, \omega_y;\ \ \omega_z, 0, -\omega_x;\ \ -\omega_y, \omega_x, 0] = sk(\omega)$ ➔ $dp/dt = \omega \times p$

General Visual Motion

Let us assume there is one rigid object moving with velocities T and ω

For a given point P on the object, we have

$$p = P/z$$

The apparent velocity of the point is

$$v = -T - \omega \times P$$

Therefore, we have

$$dp/dt = (z\, v - v_z\, P)/z^2 = 1/z\,(v - v_z\, p)$$

To Put it Another Way

$$\dot{u} = \frac{T_z u - T_x}{z} - w_y + w_z v + w_x uv - w_y u^2$$

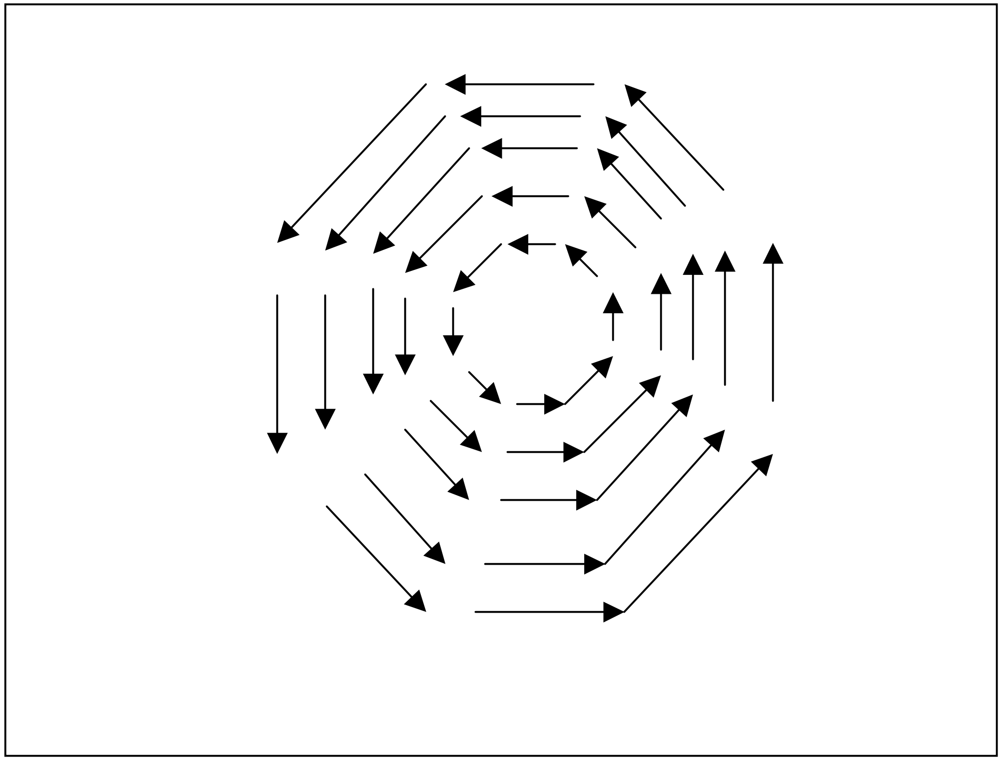
$$\dot{v} = \frac{T_z v - T_y}{z} - w_x + w_z u + w_y uv - w_x v^2$$

Motion due to translation: depends on depth

Motion due to rotation: independent of depth

THE MOTION FIELD

The “instantaneous” velocity of points in an image



PURE ROTATION

The Motion of a Planar Object

- Recall the orthographic case:
  - image transformations of planar objects are affine
  - therefore, the motion field itself is linear for planar objects
  - motion of non-planar objects can be viewed as *parallax* on top of this planar motion (note parallax only occurs to due to rotation in this model)
    - p = O s (R' P – Tr)
    - dp/dt = O s sk(ω)' P + s (T_x, T_y)' = sk_2(-ω) p + b z + d

- Now, let's generalize to the perspective case:
  - image transformations of planar objects are *homographies*
  - the motion field for planar objects is a quadratic in the coordinates of the image points
    - suggests that locally we can approximate surface motion as a quadratic
  - motion of non-planar objects can be viewed as *parallax* on top of this planar motion

Planar Optical Flow

- $n^t P = d$ --- constraint on coordinates

- Solve for P_z and do some algebra to get

$$u' = a_1x^2 + a_2xy + a_3x + a_4y + a_5$$
$$v' = a_1xy + a_2y^2 + a_7x + a_6y + a_8$$

where the a's depend on the motion and the plane parameters

It is possible to show that the motion field of a plane is not unique --- there is more than one situation that gives rise to the same apparent motion

Motion Parallax

- Consider two points on a rigid object that project to the same point instantaneously
- Note that in this case, the rotation fields are identical, hence taking a difference, we arrive at
  - $\Delta u' = (T_z u - T_x)(1/Z - 1/Z')$
  - $\Delta v' = (T_z v - T_y)(1/Z - 1/Z')$
- These can be thought of as the relative motion field
- For example, we could think of motion as a background planar motion combined with a superimposed parallax ....

$v = v_p + \Delta v$

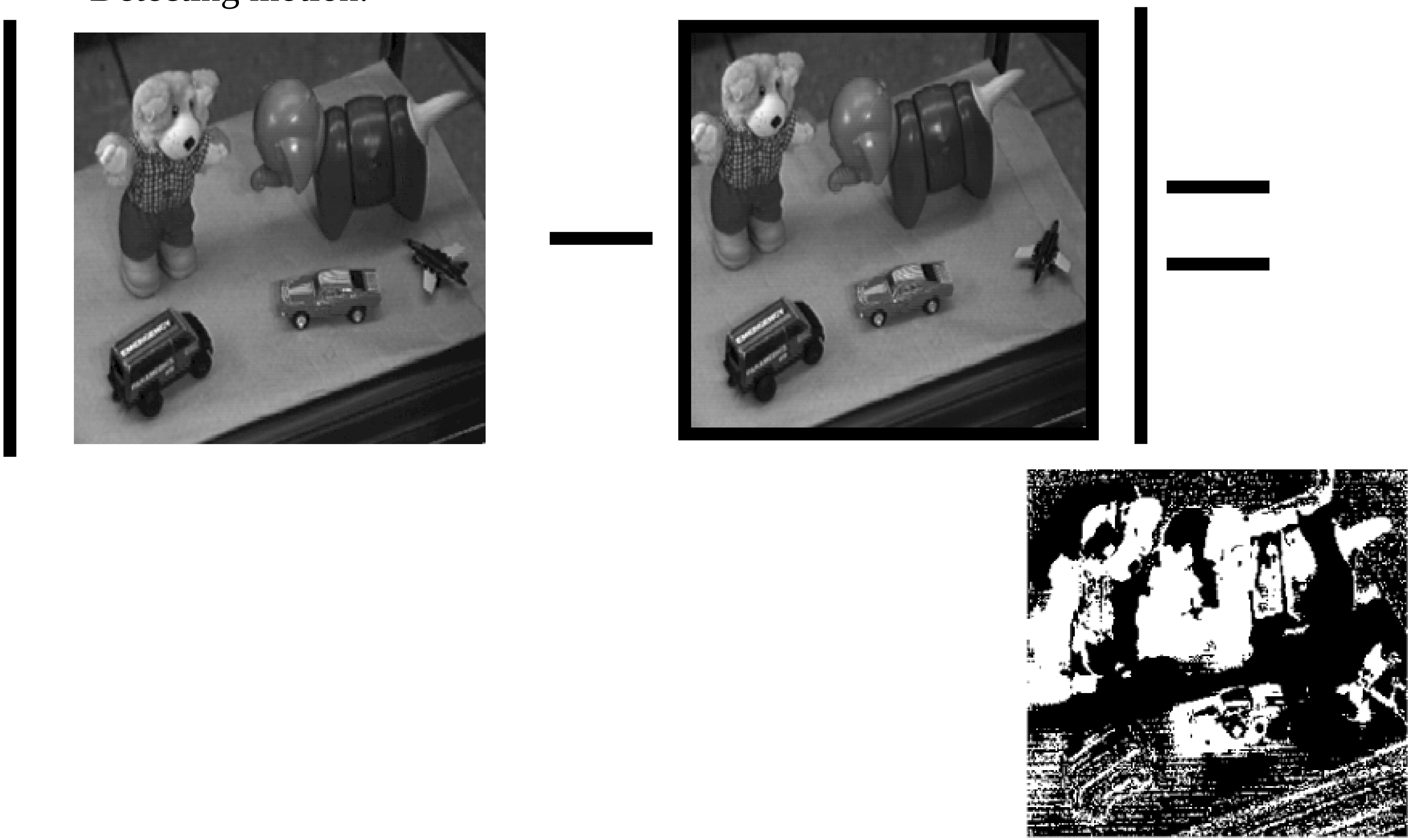
-- compute a planar motion

-- compute the effective planar motion field

-- compute Δv and estimate depth from plane

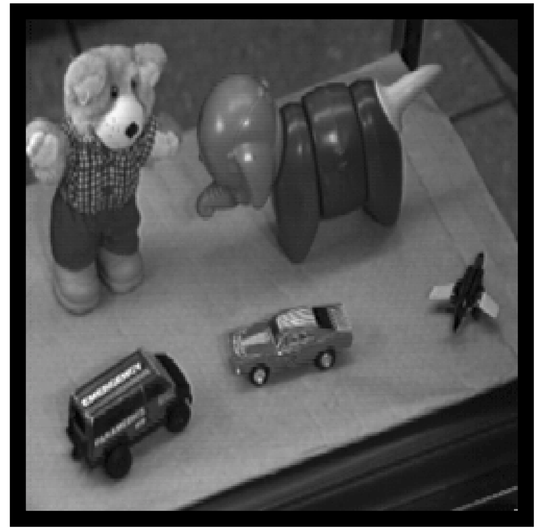
MOTION IN REAL IMAGES

Detecting motion:



MOTION IN REAL IMAGES

Detecting motion:

$\Big|$  $-$  $\Big| > 50$

Candidate areas for motion

Computing Motion: Optical Flow

- Optical flow is the *apparent motion* in the image which, in some cases, corresponds to the motion field at that point.

- To compute it, think of image brightness as a function of time:
  - $E(u(t),v(t),t)$

- Let us assume that the image brightness of a point is constant. Then we have

  - $dE/dt = dE(x(t),y(t),t)/dt = E_x\, dx/dt + E_y\, dy/dt + E_t = 0$

  - equivalently, $\nabla E \cdot v + E_t = 0$, where v is the velocity vector of an image pt.

- Note that locally, this implies that we can only compute motion perpendicular to the image gradient. This is commonly referred to as the <u>aperture problem.</u>

Computing Motion: Optical Flow

- Since we can’t solve the IC equation at a point, two general approaches to modifying the problem:
  - regularization
  - finite patch
- Regularization assumes that points move locally consistently
  - $O(v(i)) = \nabla E(i) . v(i) + E_t(i) + \sum_j || v(j) – v(i) ||^2$
  - solution is given by a linear system over the entire image!
  - local iterations approach solution

Computing Optical Flow

- Finite patch model:
  - $O(v(i)) = \sum_j || \nabla E(j) . v(i) + E_t(j)||^2$
  - $O(v(i)) = || A v(i) + b ||^2$

    where $A = [\nabla E(1); \nabla E(2); ... \nabla E(n)]$ $\quad b = [E_t(1); E_t(2); ... E_t(n)]$

  - $O'(v(i)) = A' (A v(i) + b) = 0 \rightarrow v(i) = - (A'A)^{-1} A' b$
- Algorithm:
  - Filter in u,v, and t using appropriate smoothing derivative
  - Solve previous equations for each point in image

Remarks

- When is A'A full rank?
  - when spatial gradient spans R^2
  - SVD of A gives a measure of confidence
- How fast can we track?
  - derivatives are good for ½ pixel
  - smoothing correlates pixels and increases attraction
  - subsampling decreases # of pixels and increases effective pixel size
  - → usually Oflow is done in a heirarchical computation.

When is the ICC Valid?

- ..... the same development as in the book ...