6.801/866

Optic Flow and Direct Motion

T. Darrell

Material kindly adapted from B. Horn, Y. Weiss, P. Anandan, M. Black, K. Toyama

Topics

- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- Direct depth
- SSD tracking
- Robust flow
- Bayesian flow

Applications

- tracking
- recognition
- structure from motion
- segmentation
- stabilization
- compression
- mosaicing

...

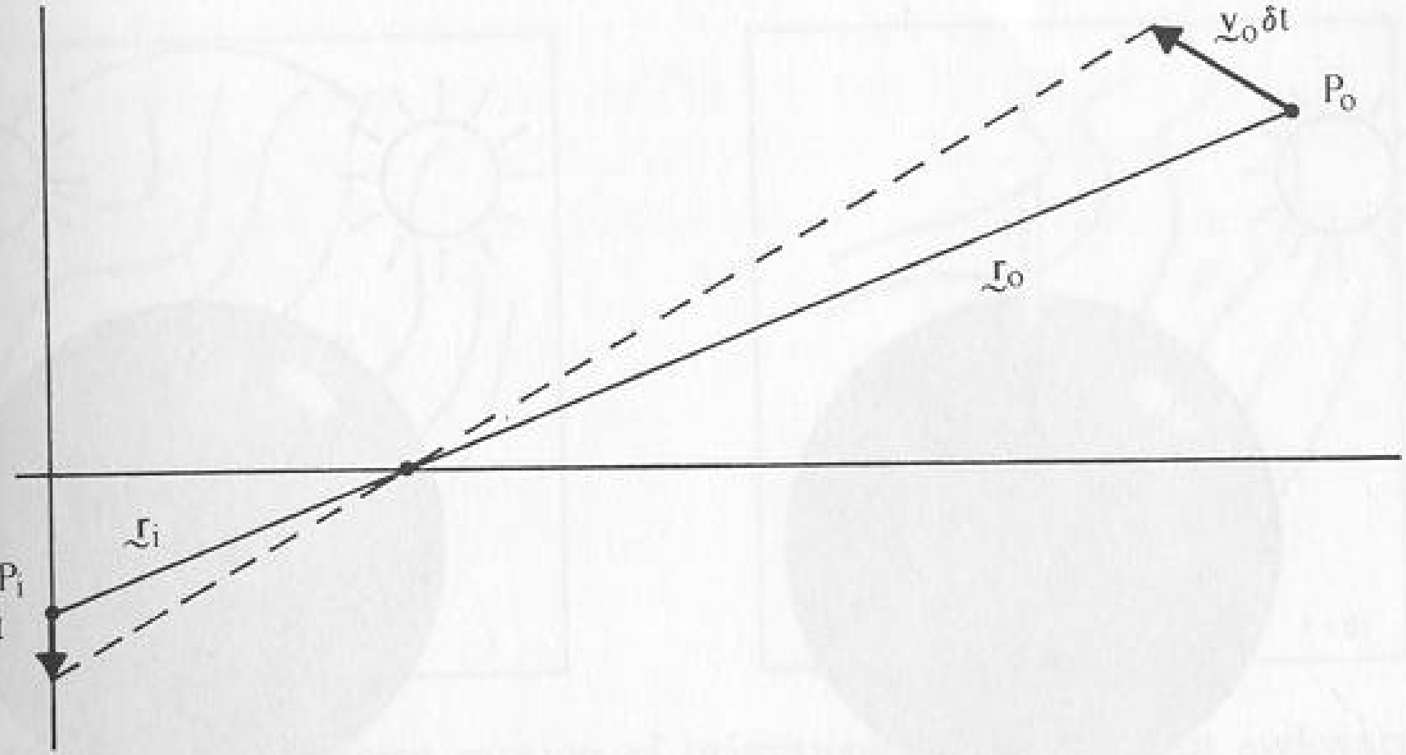


Figure 12-1. Displacement of a point in the environment causes a displacement of the corresponding image point. The relationship between the velocities can be found by differentiating the perspective projection equation.

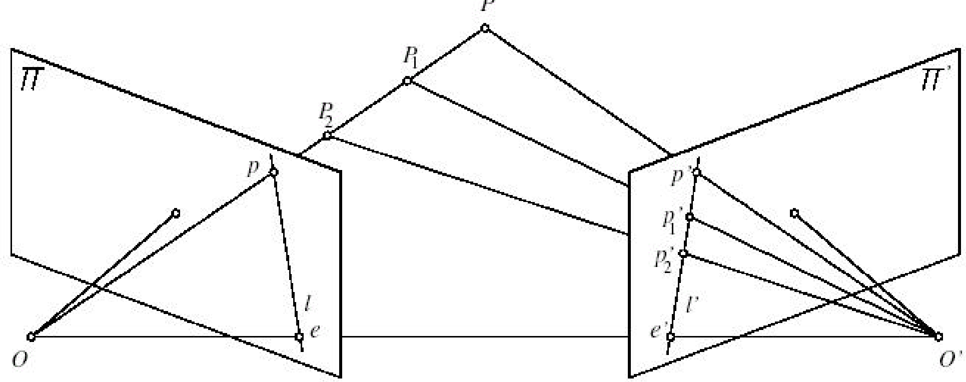
Review: The Essential Matrix

Matrix that relates image of point in one camera to a second camera, given translation and rotation.

5 independent parameters (up to scale)

Assumes intrinsic parameters are known.

$$e = [t_x]\Re$$



$$\boldsymbol{p}^T \mathcal{E} \boldsymbol{p}' = 0$$

$$\vec{a} \times \vec{b} = [a_x]\vec{b}$$

Review: Instantaneous case

For small motion given translation and rotation velocity:

$$\boldsymbol{t} = \delta t\, \boldsymbol{v},$$
$$\mathcal{R} = \mathrm{Id} + \delta t\, [\boldsymbol{\omega}_{\times}]$$
$$\boldsymbol{p}' = \boldsymbol{p} + \delta t\, \dot{\boldsymbol{p}}.$$

$$\boldsymbol{p}^T [\boldsymbol{v}_{\times}] (\mathrm{Id} + \delta t\, [\boldsymbol{\omega}_{\times}]) (\boldsymbol{p} + \delta t\, \dot{\boldsymbol{p}}) = 0$$

See F&P, chapter 10

Review: FOE for translating camera

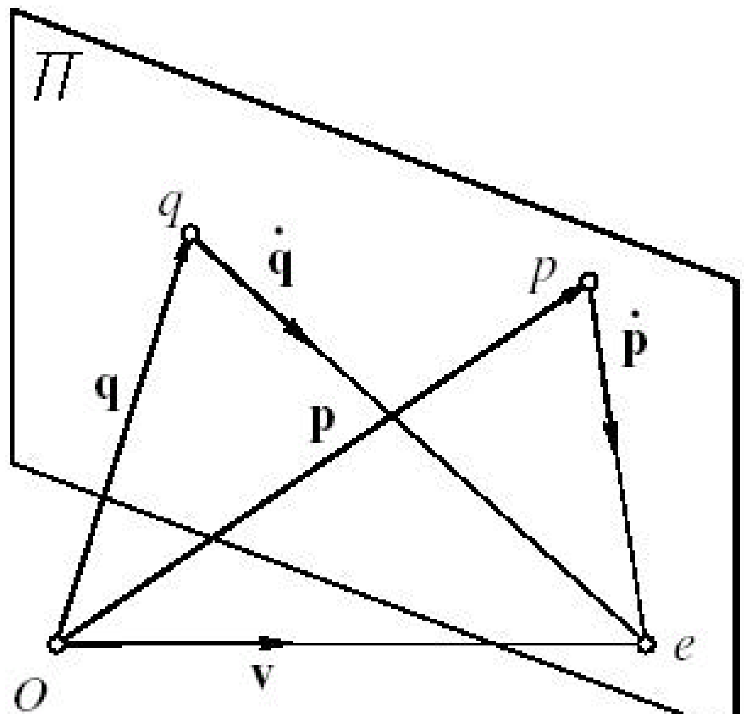
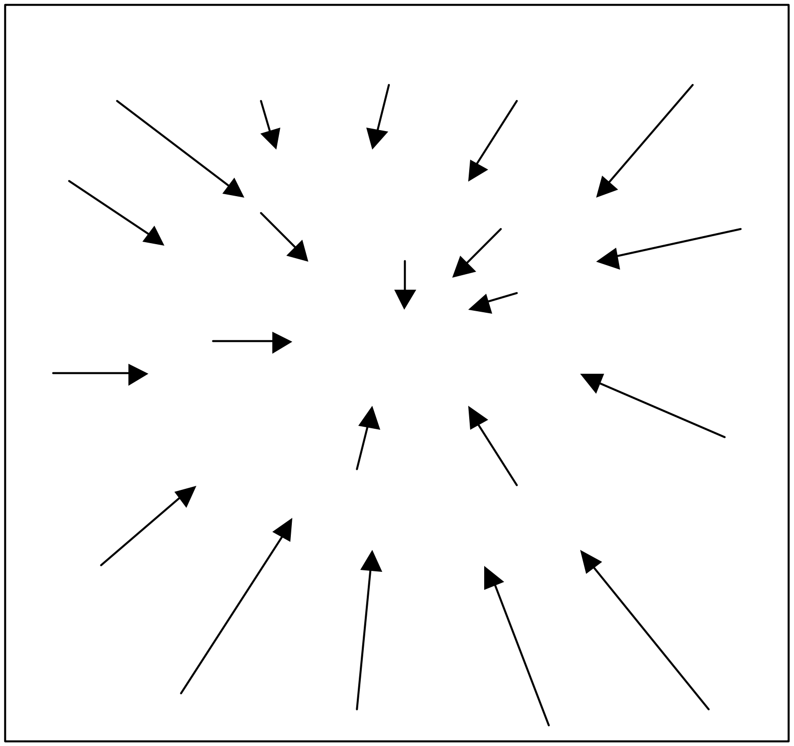


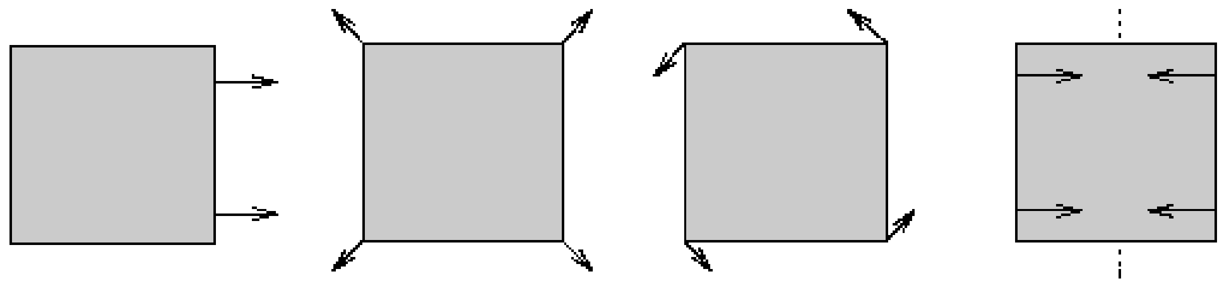
FIGURE 11.3: Focus of expansion: under pure translation, the motion field at every point in the image points toward the focus of expansion.

Review: FOE for translating camera



Planar motion examples

- Ideal motion of a plane



translation in X

translation in Z

rotation around Z

rotation around Y

How to find motion?

- Track features (as in stereo)?
  - exhaustive search can be slow
  - usually just return best match; no uncertainty
- Differential approach
  - match gradients
  - usually limited by first-order Taylor series approximation
  - most common in CV literature

Differential approach:
Optical flow constraint equation

Brightness should stay constant as you track motion

$$I(x+udt, y+vdt, t+dt) = I(x, y, t)$$

1st order Taylor series, valid for small dt

$$I(x, y, t) + udtI_x + vdtI_y + dtI_t = I(x, y, t)$$

Constraint equation

$$\boxed{uI_x + vI_y + I_t = 0}$$

"BCCE" - Brightness Change Constraint Equation

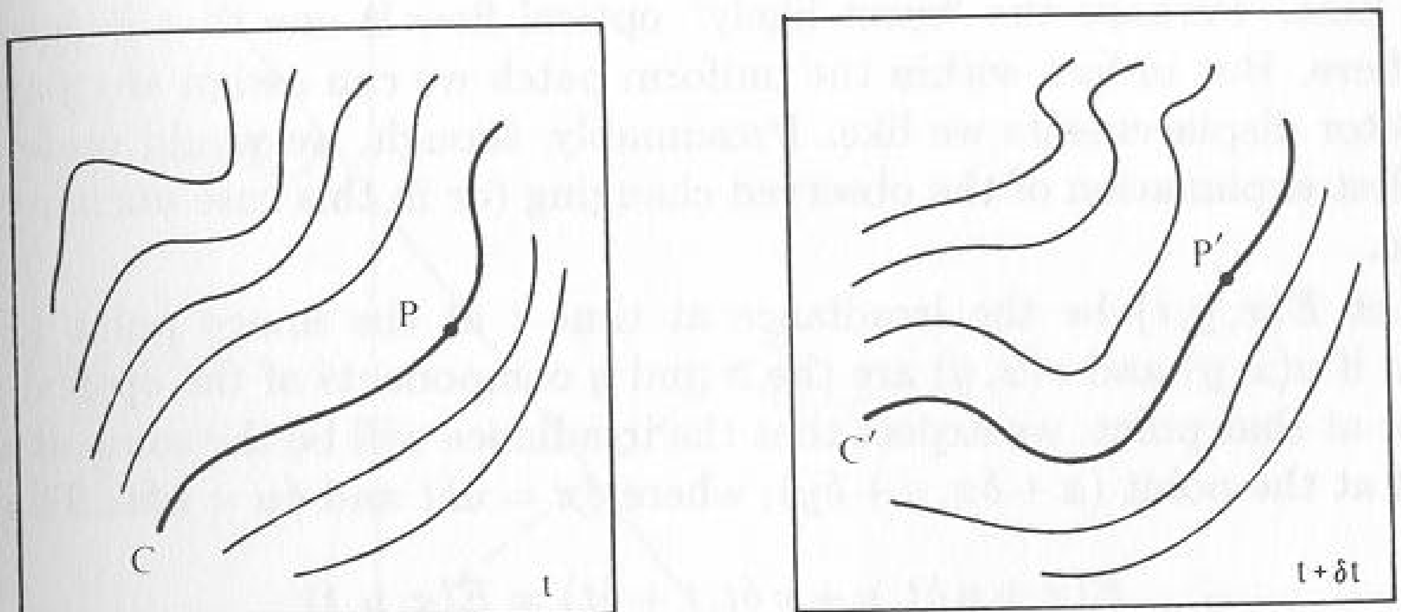
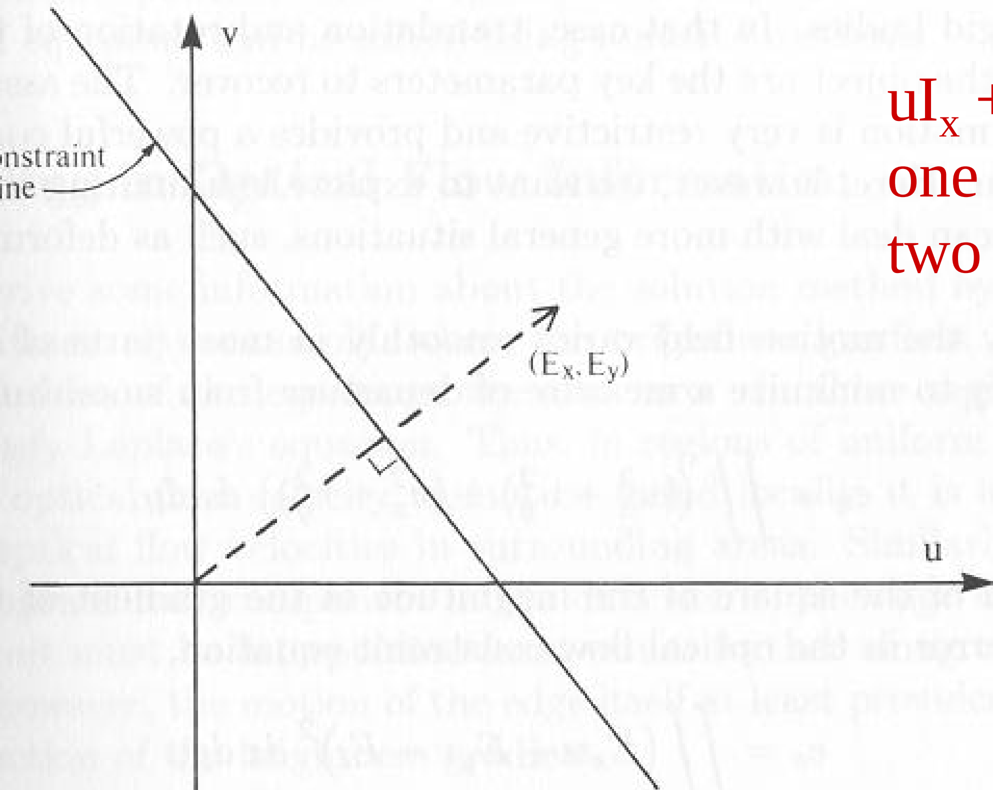


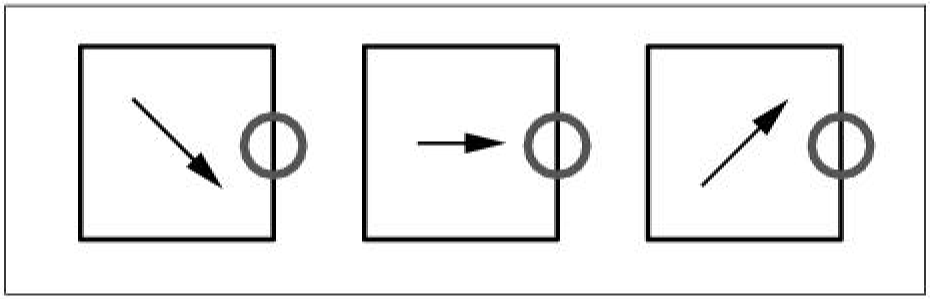
Figure 12-3. The apparent motion of brightness patterns is an awkward concept. It is not easy to decide which point P' on a contour C' of constant brightness in the second image corresponds to a particular point P on the corresponding contour C in the first image.

Brightness constancy constraint line

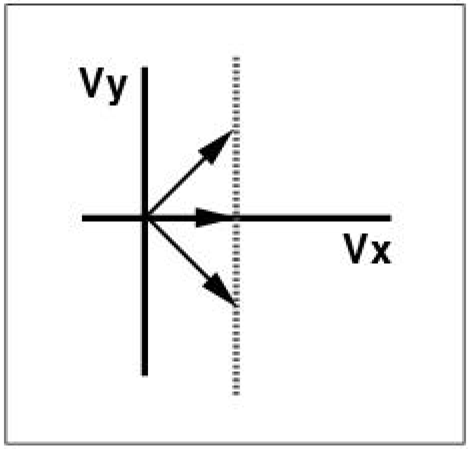


$uI_x + vI_y + I_t = 0$:
one equation,
two unknowns!

Figure 12-4. Local information on the brightness gradient and the rate of change of brightness with time provides only one constraint on the components of the optical flow vector. The flow velocity has to lie along a straight line perpendicular to the direction of the brightness gradient. We can only determine the component in the direction of the brightness gradient. Nothing is known about the flow component in the direction at right angles.



a

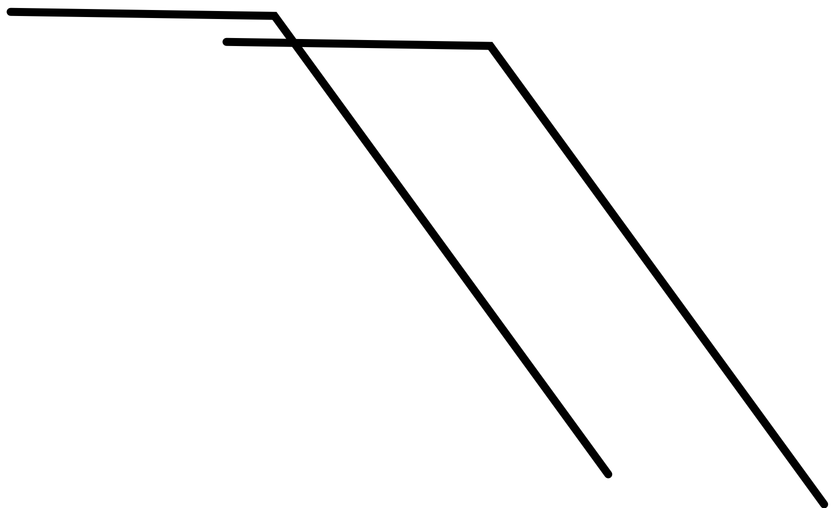


b

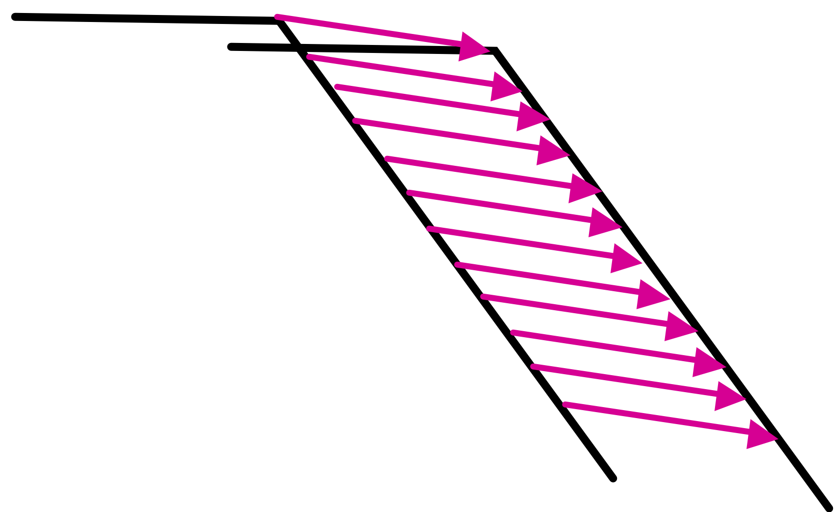
Weiss and Fleet, in Rao and Sejnowski (ed) Statistical Theories of the Cortex. MIT Press

Figure 1: **a.** The "aperture problem" refers to the inability to determine the two dimensional motion of a signal containing a single orientation. For example, a local analyzer that sees only the vertical edge of a square can only determine the horizontal component of the motion. Whether the square translates horizontally to the right, diagonally up and to the right, or diagonally down and to the right, the motion of the vertical edge will be the same. **b.** The family of motions consistent with the motion of the edge can be depicted as a line in "velocity space", where any velocity is represented as a vector from the origin whose length is proportional to speed and whose angle corresponds to direction of motion. Graphically, the aperture problem is equivalent to saying that the family of motions consistent with the edge maps to a straight line in velocity space, rather than a single point.

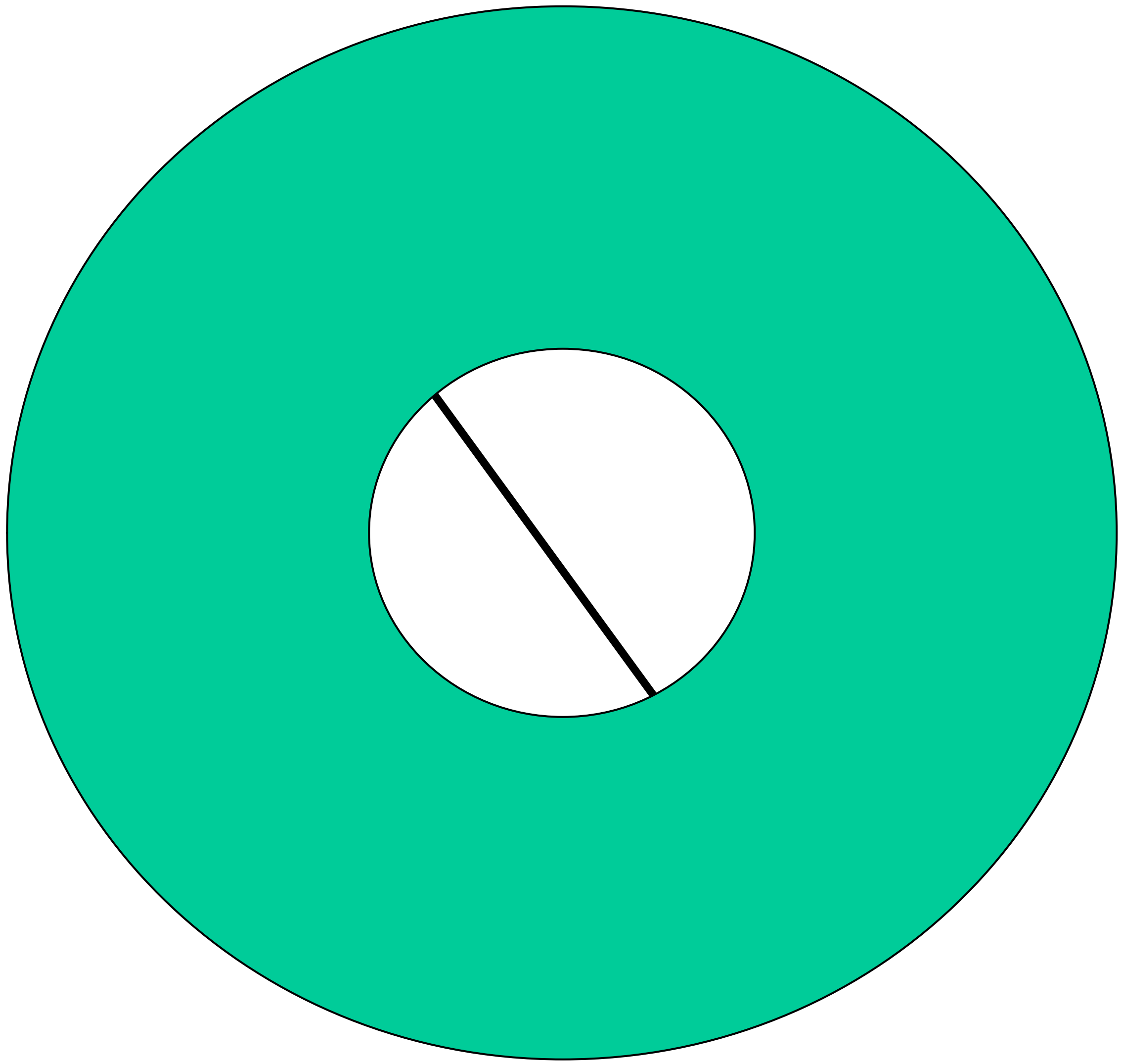
Aperture Problem and Normal Flow



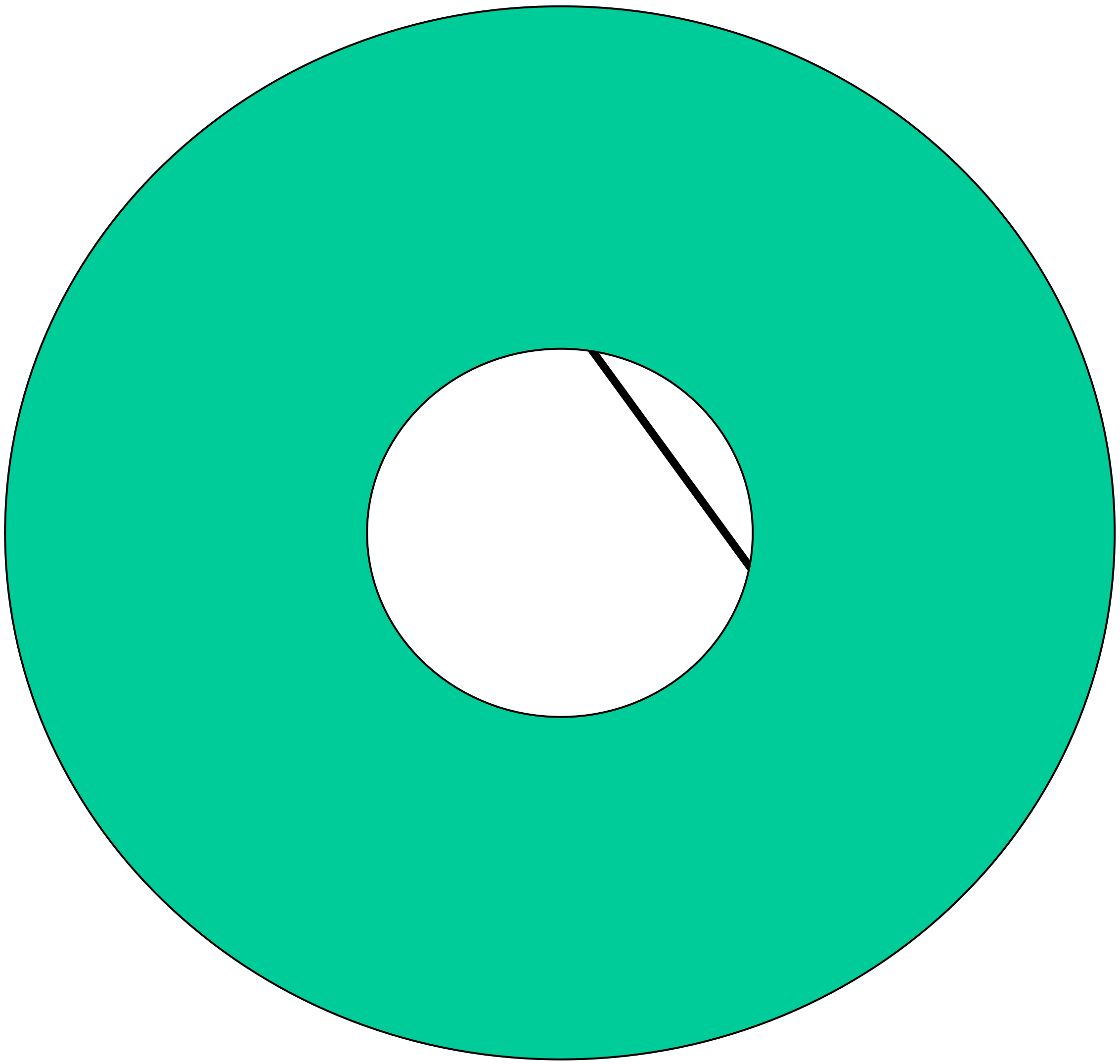
Aperture Problem and Normal Flow



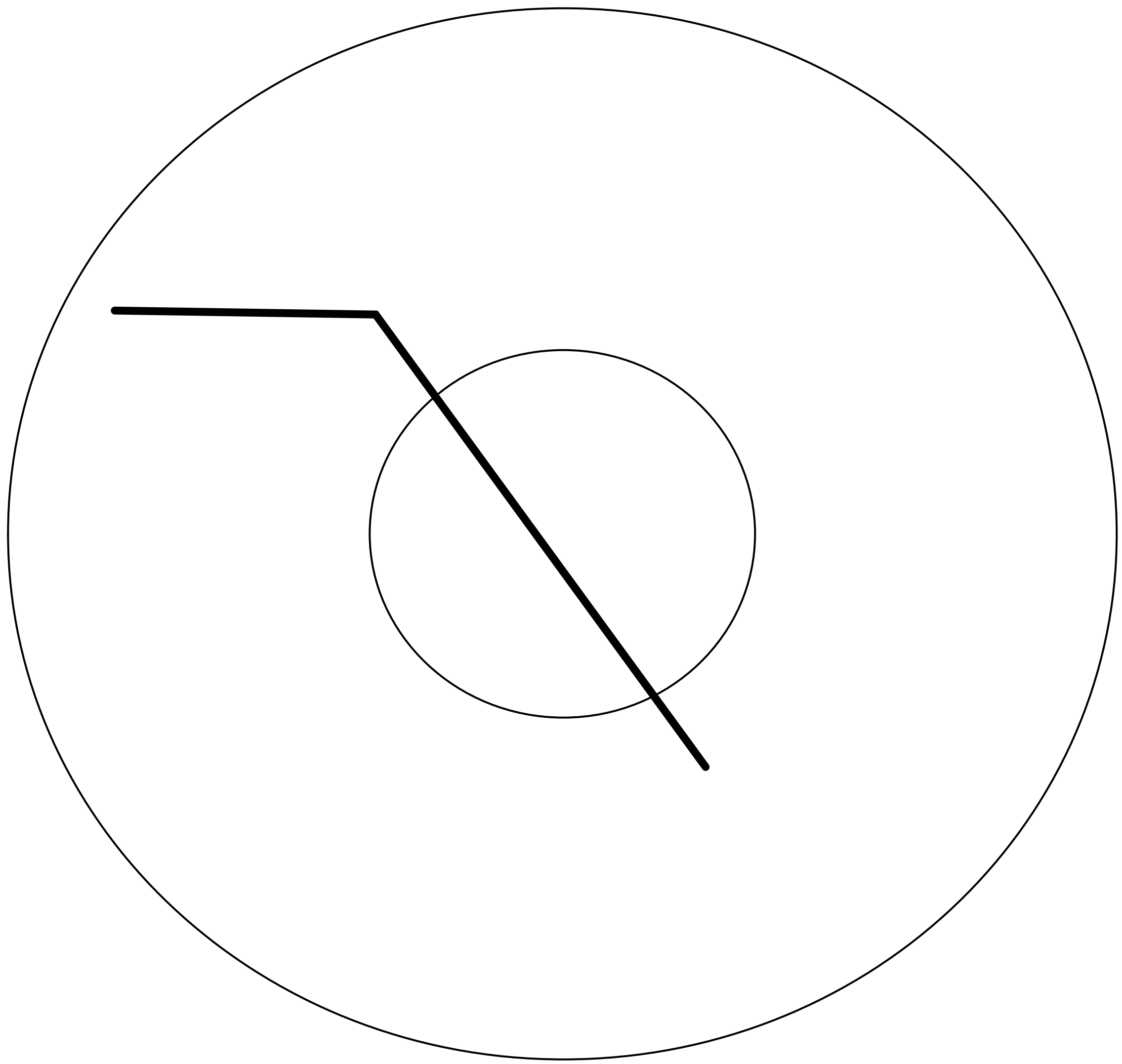
Aperture Problem and Normal Flow



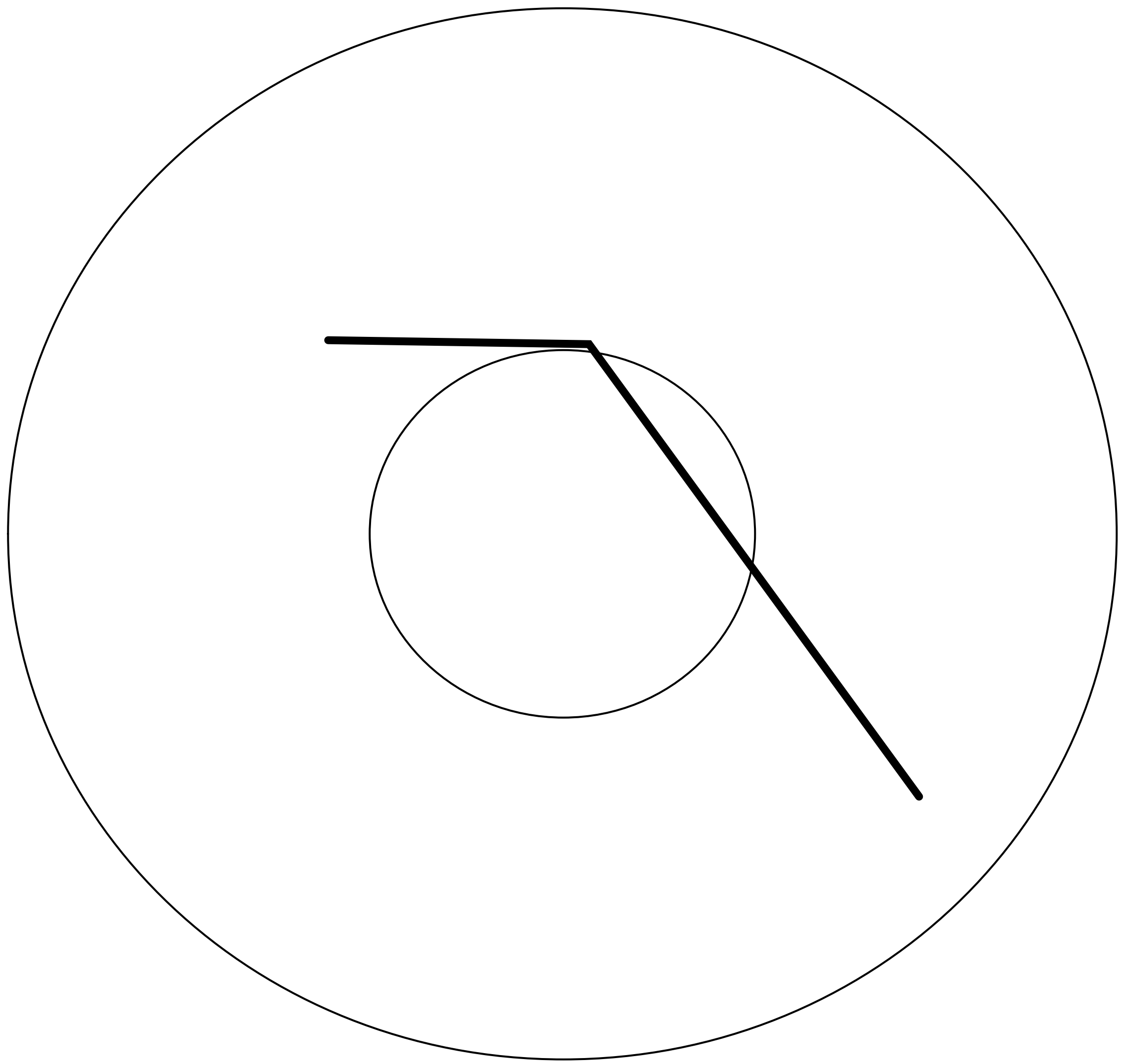
Aperture Problem and Normal Flow



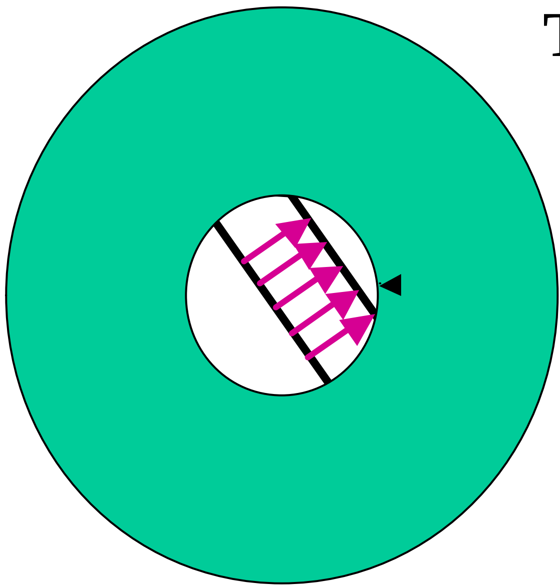
Aperture Problem and Normal Flow



Aperture Problem and Normal Flow



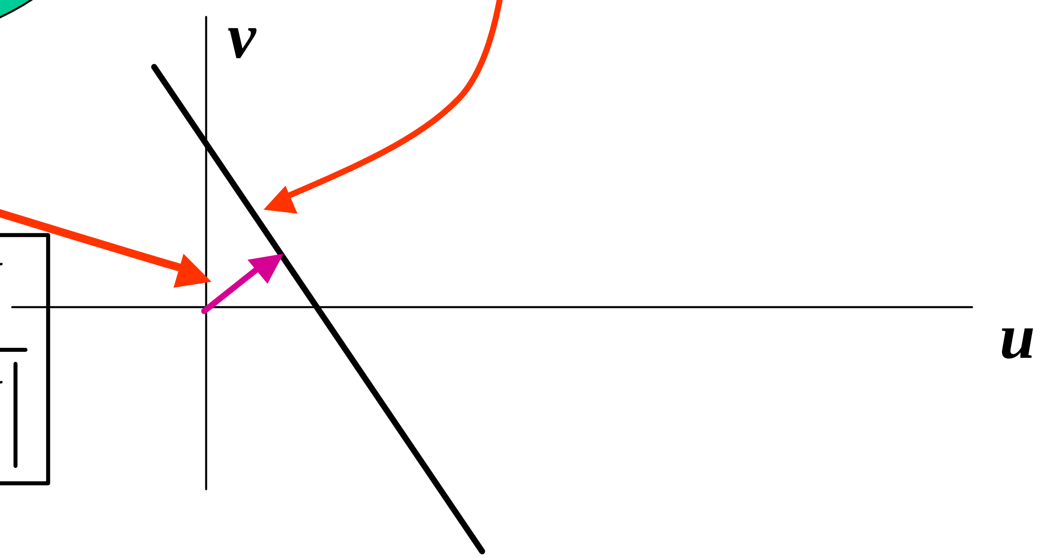
Aperture Problem and Normal Flow



The gradient constraint:

$$\boxed{\begin{aligned} I_x u + I_y v + I_t &= 0 \\ \nabla I \bullet \vec{U} &= 0 \end{aligned}}$$

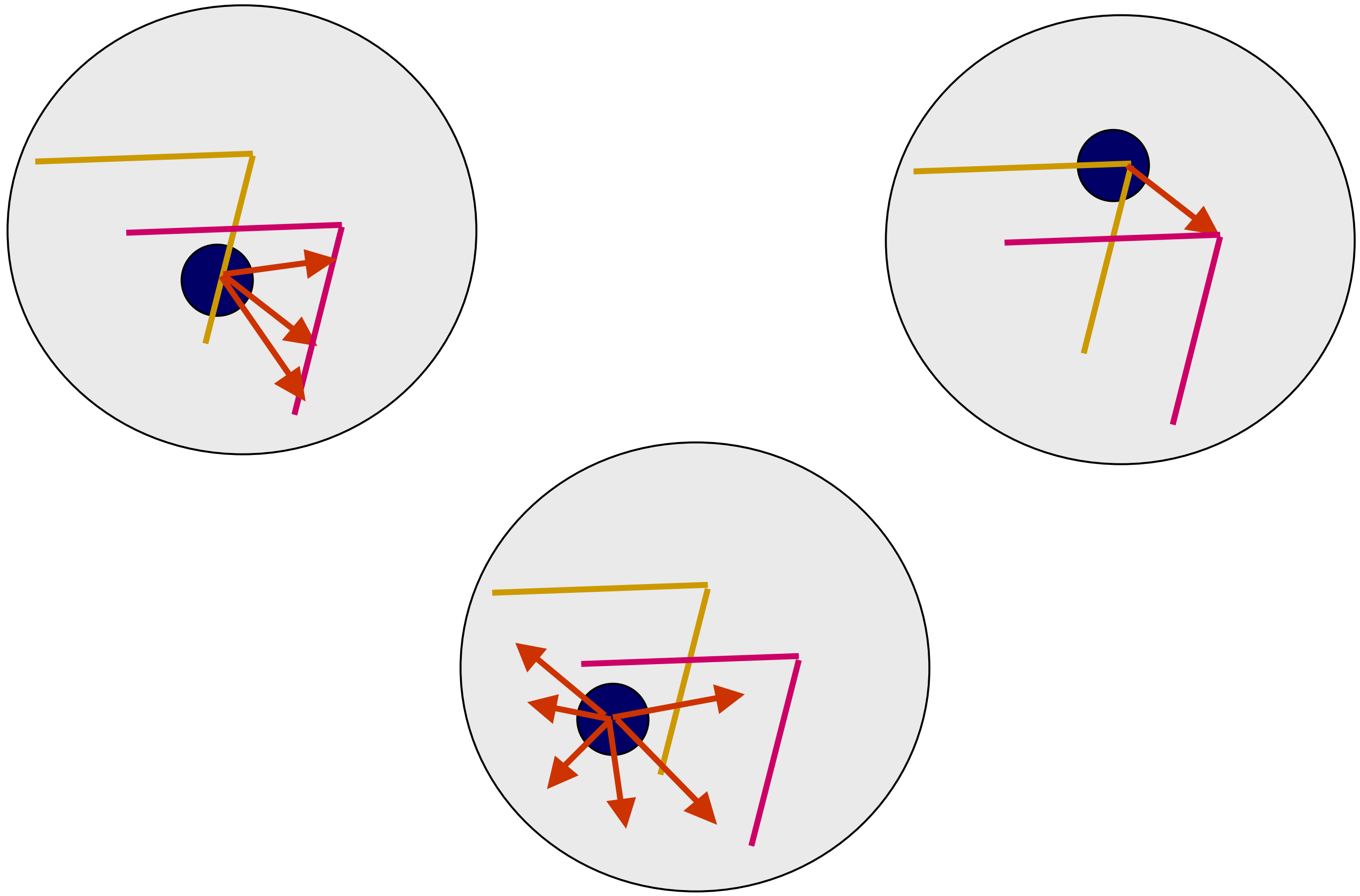
Defines a line in the *(u,v)* space



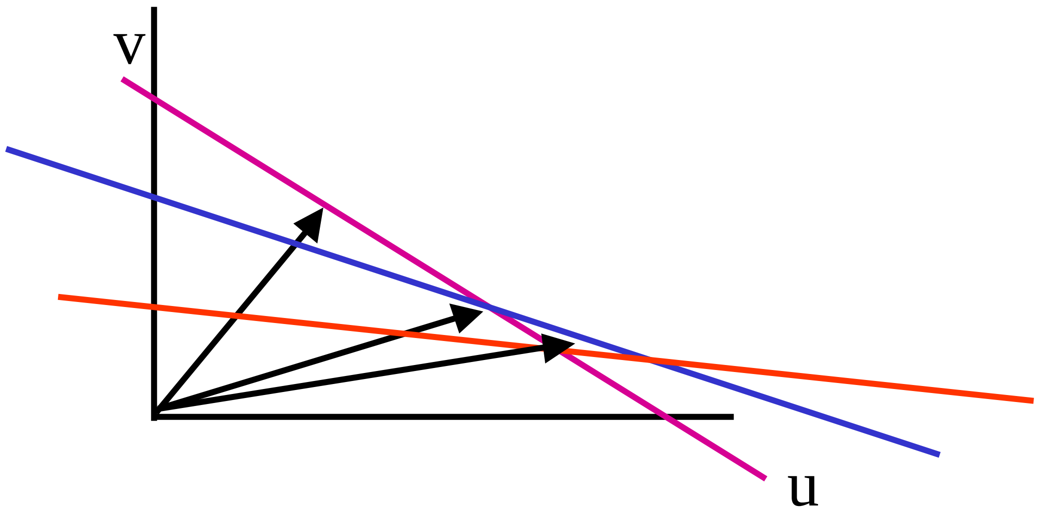
Normal Flow:

$$\boxed{u_\perp = -\frac{I_t}{|\nabla I|}\frac{\nabla I}{|\nabla I|}}$$

Local Patch Analysis



Combining Local Constraints



$$\nabla I^1 \bullet U = -I_t^1$$

$$\nabla I^2 \bullet U = -I_t^2$$

$$\nabla I^3 \bullet U = -I_t^3$$

etc.

What is Optic Flow, anyway?

- Estimate of observed projected motion field
- Not always well defined!
- Compare:
  - Motion Field (or Scene Flow)

    projection of 3-D motion field
  - Normal Flow

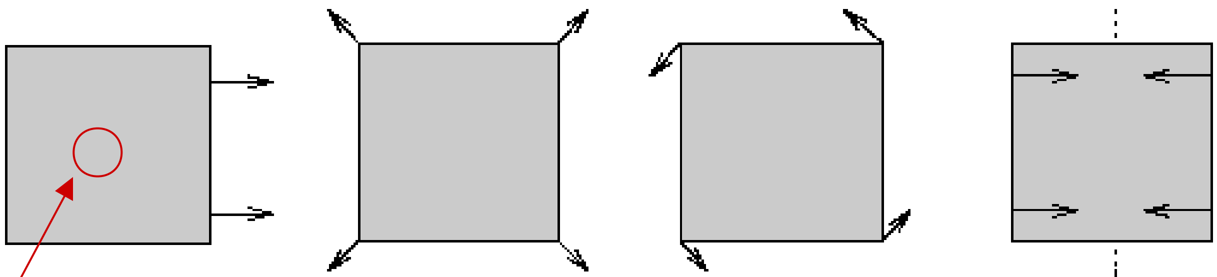
    observed tangent motion
  - Optic Flow

    apparent motion of the brightness pattern

    (hopefully equal to motion field)
- Consider Barber pole illusion

Planar motion examples

- Ideal motion of a plane



What is the motion here?

translation in X

translation in Z

rotation around Z

rotation around Y

Scene Flow: →

Normal Flow: undef

Optic Flow: ?, probably 0

Topics

- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- Direct depth
- SSD tracking
- Robust flow
- Bayesian flow

Solutions to “the aperture problem”

- Regularize the solution: add a velocity smoothness constraint (eg, Horn 12.6).
- Integrate over a larger region in image (e.g., Lukas and Kanade).
- More sophisticated scene models: segment object boundaries, specularities, textureless regions, etc.

Horn and Schunck

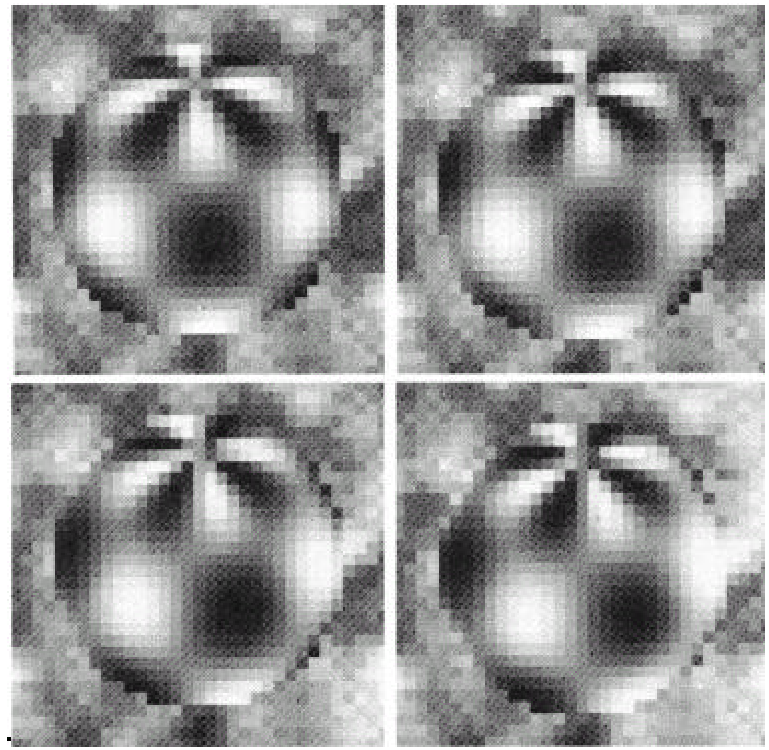
- Smoothness is most natural:
  - find **v** that minimizes

$$\iint \left(uI_x + vI_y + I_t\right)^2 + \alpha^2\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right) dx\,dy$$

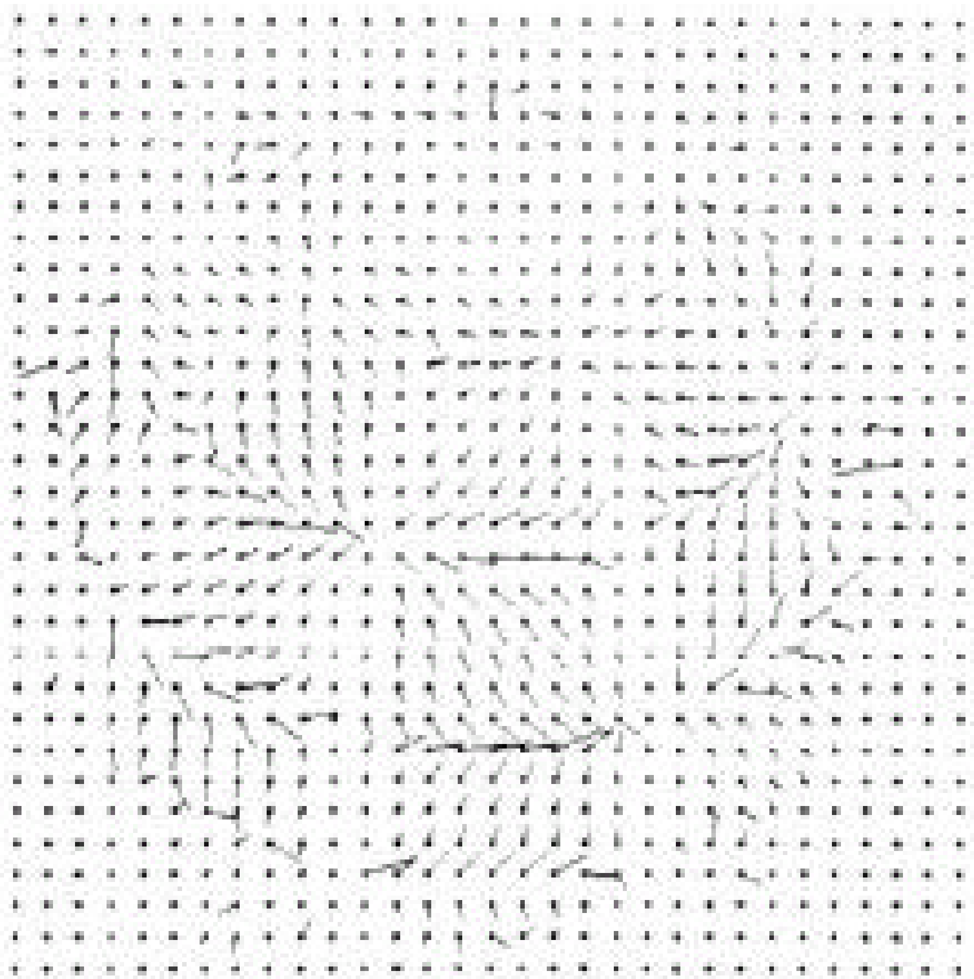
$$I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}$$

$$u = \frac{dx}{dt}, v = \frac{dy}{dt}, u_x = \frac{\partial u}{\partial x}, v_x = \frac{\partial v}{\partial x}, u_y = \frac{\partial u}{\partial y}, v_y = \frac{\partial v}{\partial y}$$

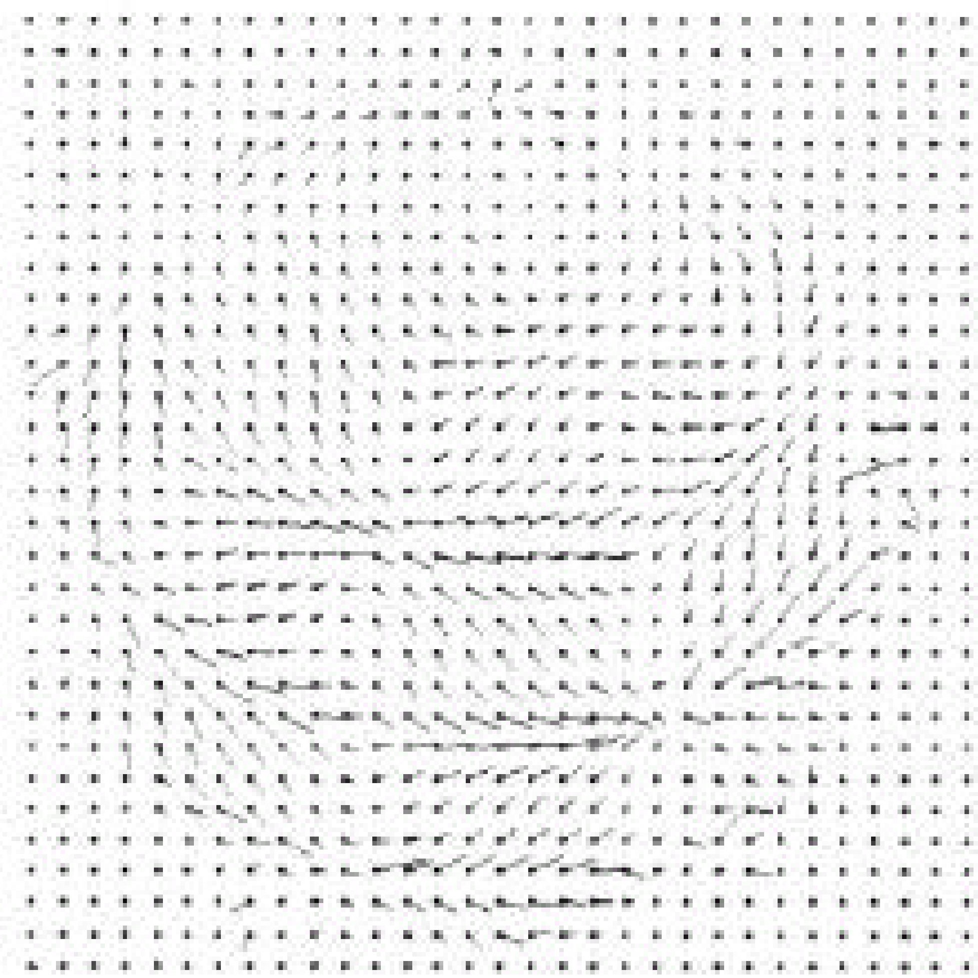
Horn and Schunck



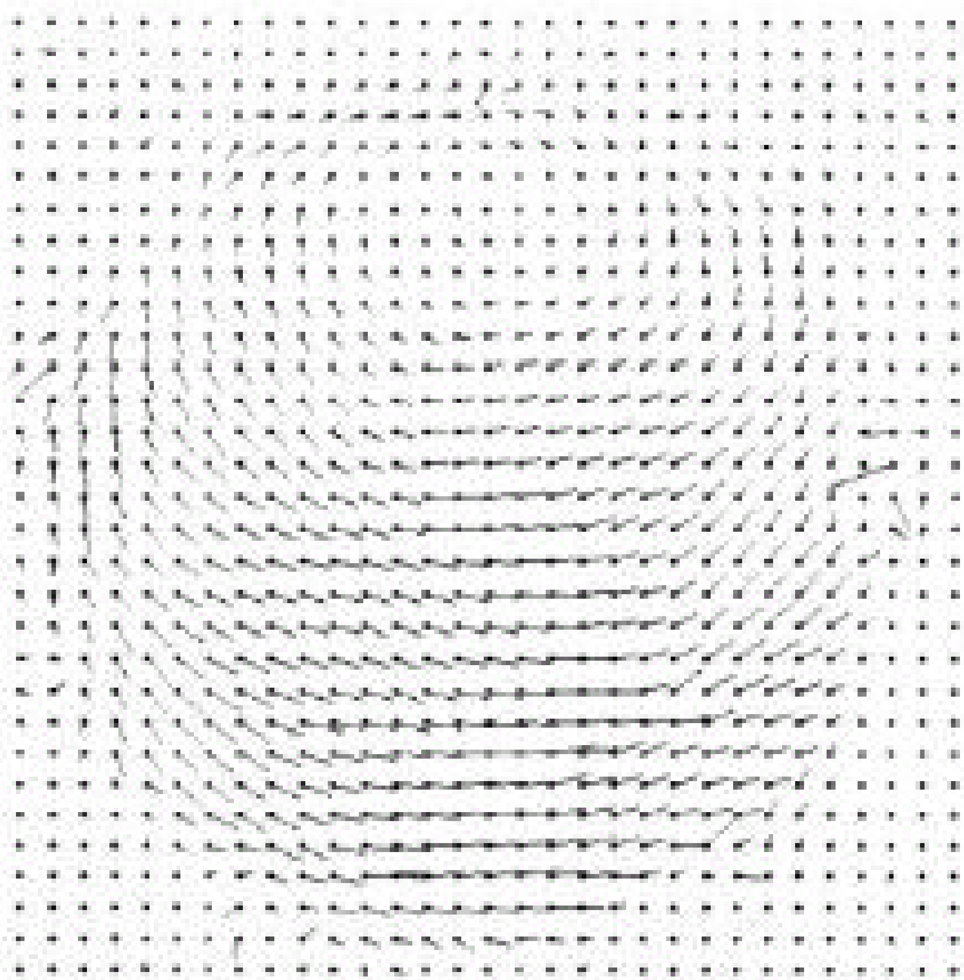
Horn and Schunck



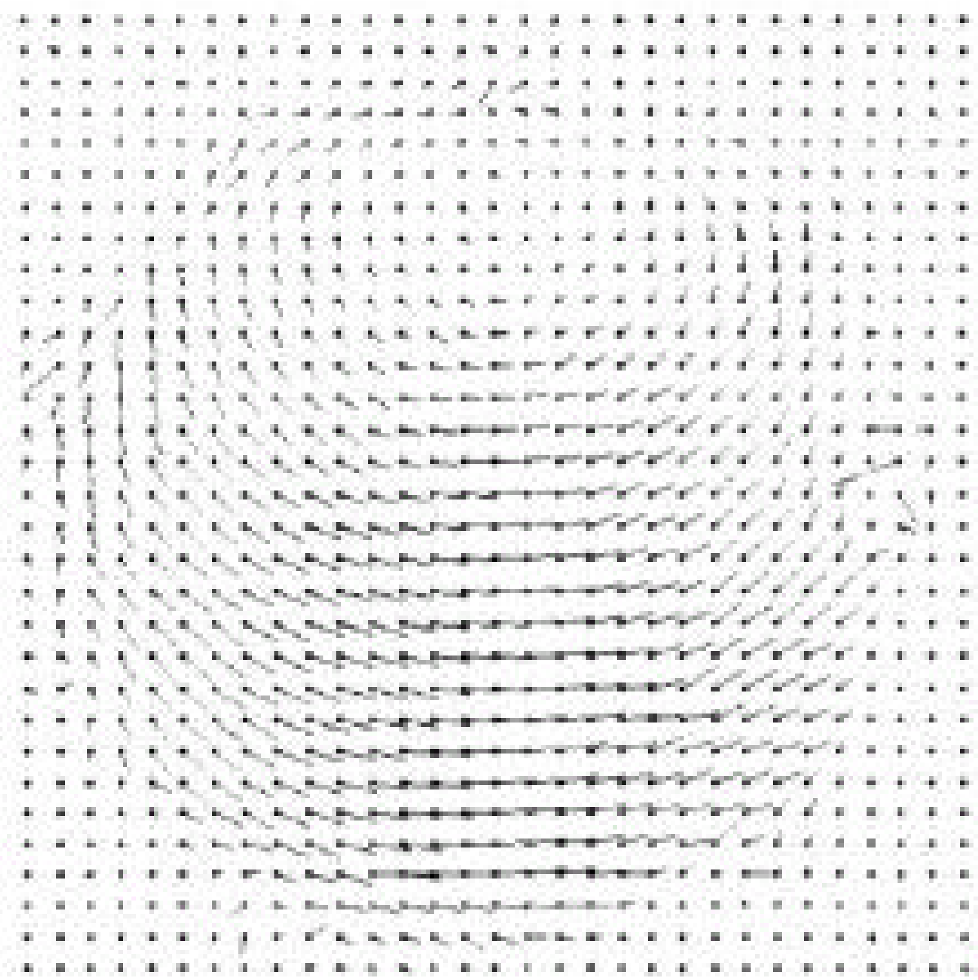
Horn and Schunck

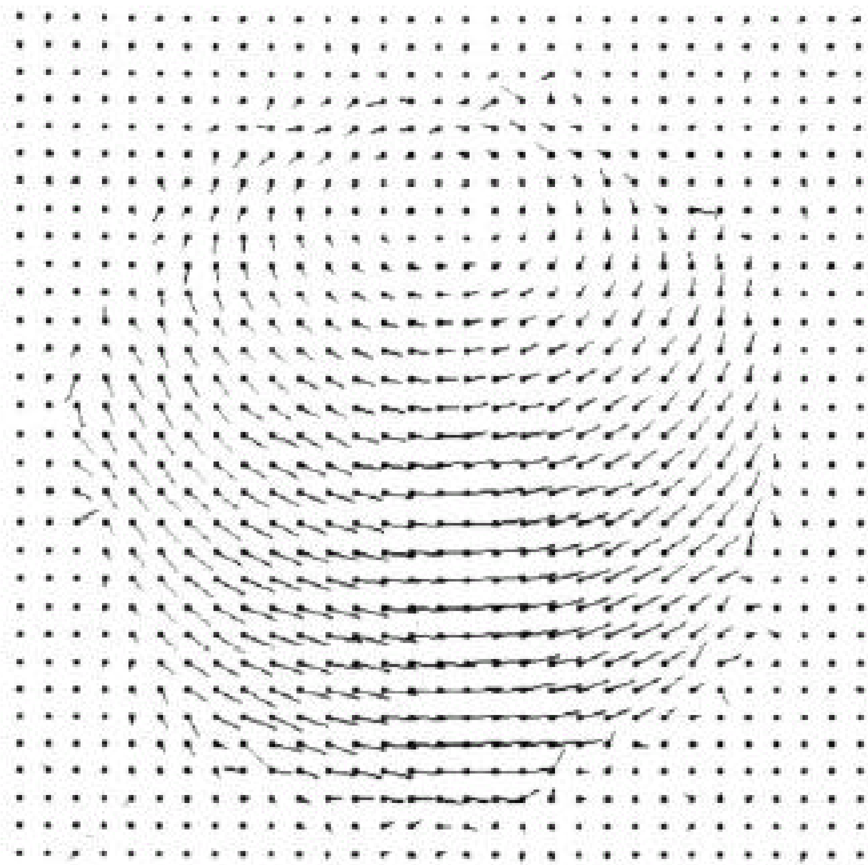


Horn and Schunck

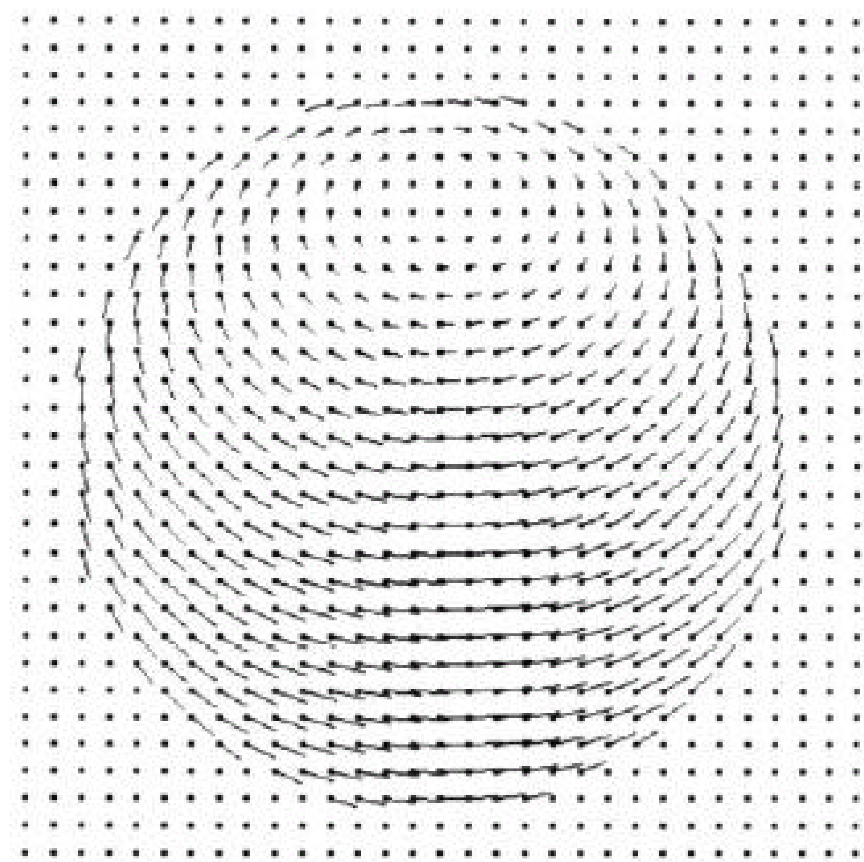


Horn and Schunck





Final estimate



Actual motion

Problems:

–edges

–large uniform regions

–can require many iterations

Lucas-Kanade: Integrate over a Patch

Assume a single velocity for all pixels within an image patch

$$E(u,v) = \sum_{x,y\in\Omega} \left(I_x(x,y)u + I_y(x,y)v + I_t \right)^2$$

Solve with:

$$\begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = -\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}$$

On the LHS: sum of the 2x2 outer product tensor of the gradient vector

$$\left(\sum \nabla I \nabla I^T \right) \vec{U} = -\sum \nabla I I_t$$

Lucas-Kanade: Singularities and the Aperture Problem

Let $M = \sum (\nabla I)(\nabla I)^T$ and $b = \begin{bmatrix} -\sum I_x I_t \\ -\sum I_y I_t \end{bmatrix}$

- Algorithm: At each pixel compute U by solving $MU = b$

- M is singular if all gradient vectors point in the same direction
  - -- e.g., along an edge
  - -- of course, trivially singular if the summation is over a single pixel
  - -- i.e., only *normal flow* is available (aperture problem)

- Corners and textured areas are OK

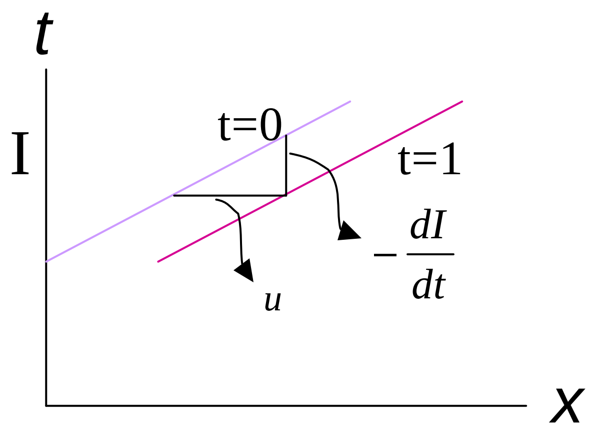
Iterative Refinement

- Estimate velocity at each pixel using one iteration of Lucas and Kanade estimation
- Warp one image toward the other using the estimated flow field

  (easier said than done)
- Refine estimate by repeating the process

Motion and Gradients

Consider 1-d signal; assume linear function of x



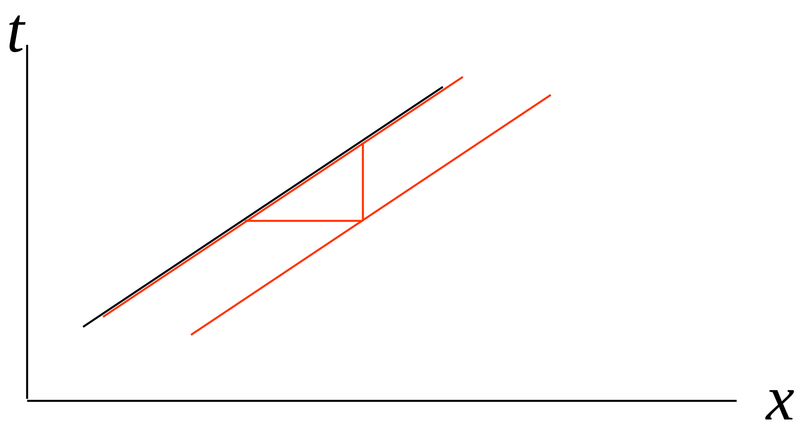
$$\frac{dI}{dx} = -\frac{dI/dt}{u}$$

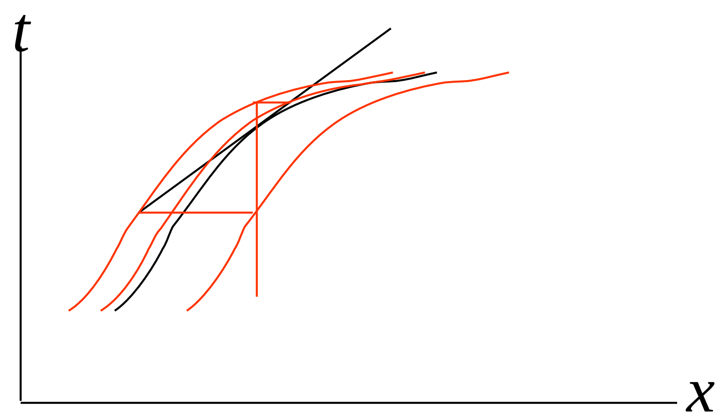
$$0 = I_x u + I_t$$

"shift by u to account for I_x with I_t"

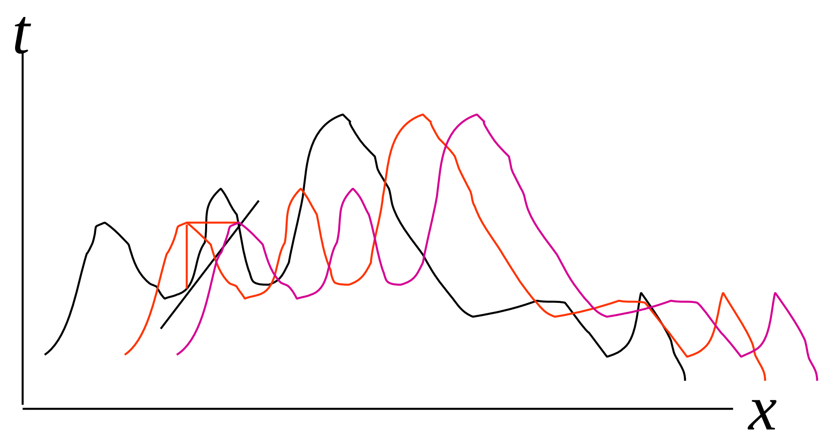
$$u = -\frac{I_t}{I_x}$$

Iterative refinement





BUT!!



Topics

- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- Direct depth
- SSD tracking
- Robust flow
- Bayesian flow

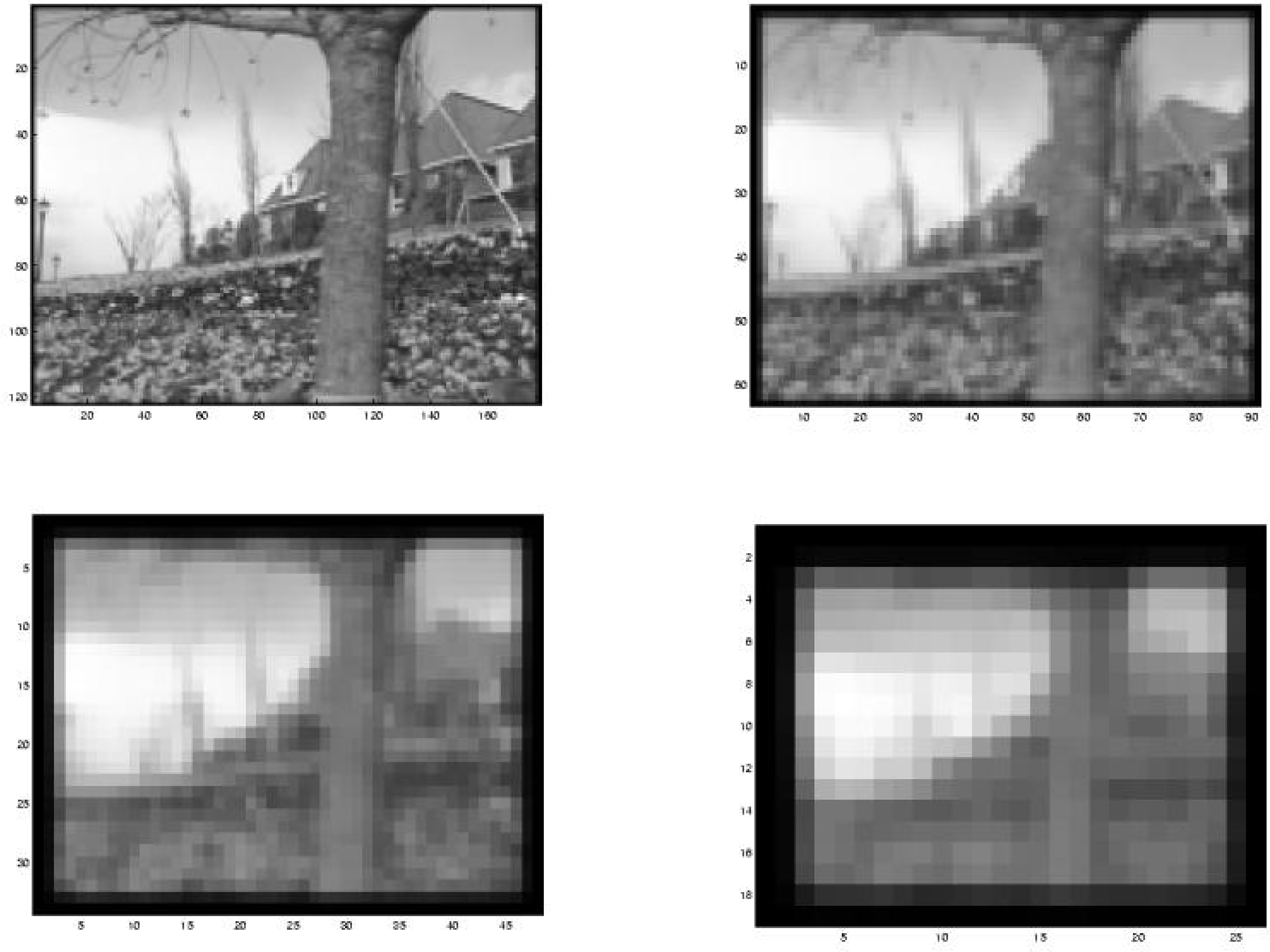
Limits of the (local) gradient method

1. Fails when intensity structure within window is poor
2. Fails when the displacement is large (typical operating range is motion of 1 pixel per iteration!)
   - *Linearization of brightness is suitable only for small displacements*

Also, brightness is not strictly constant in images

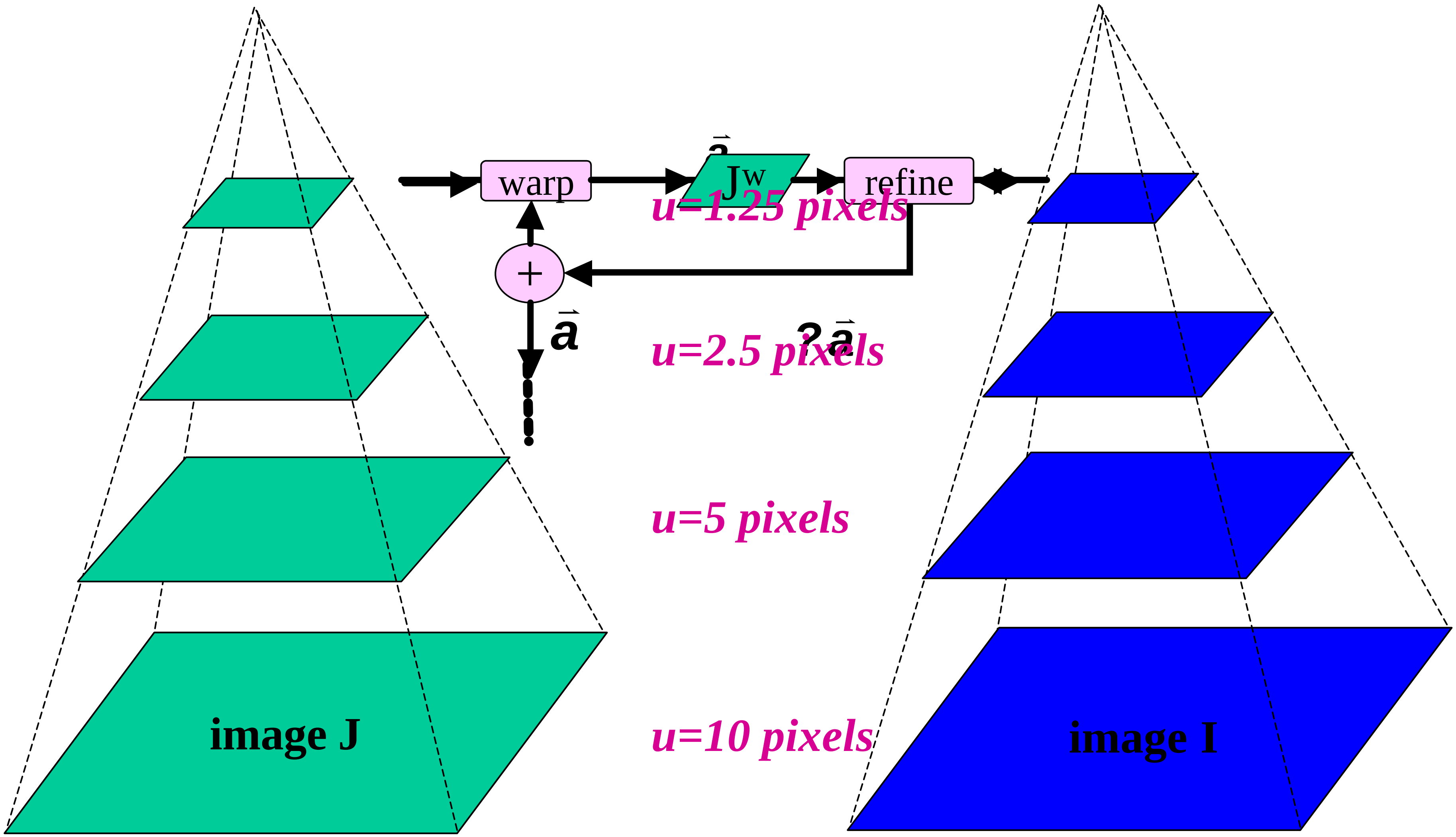
- *actually less problematic than it appears, since we can pre-filter images to make them look similar*

Pyramid / “Coarse-to-fine”



Coarse-to-Fine Estimation

$$\boxed{I_x \cdot u + I_y \cdot v + I_t \approx 0}$$ **==> small *u* and *v* ...**



Pyramid of image J **Pyramid of image I**

Parametric (Global) Motion Models

Global motion models offer

- more constrained solutions than smoothness (Horn-Schunck)
- integration over a larger area than a translation-only model can accommodate (Lucas-Kanade)

Parametric (Global) Motion Models

2D Models:

(Translation)

Affine

Quadratic

Planar projective transform (Homography)

3D Models:

Instantaneous camera motion models

Homography+epipole

Plane+Parallax

$$E(\mathbf{h}) = \sum_{\mathbf{x} \in R} \left[I(\mathbf{x} + \mathbf{h}) - I_0(\mathbf{x}) \right]^2$$

- Transformations/warping of image

$$E(\mathbf{h}) = \sum_{\mathbf{x} \in R} \left[I(\mathbf{x} + \mathbf{h}) - I_0(\mathbf{x}) \right]^2$$

Translations:

$$\mathbf{h} = \begin{bmatrix} dx \\ dy \end{bmatrix}$$

What about other types of motion?

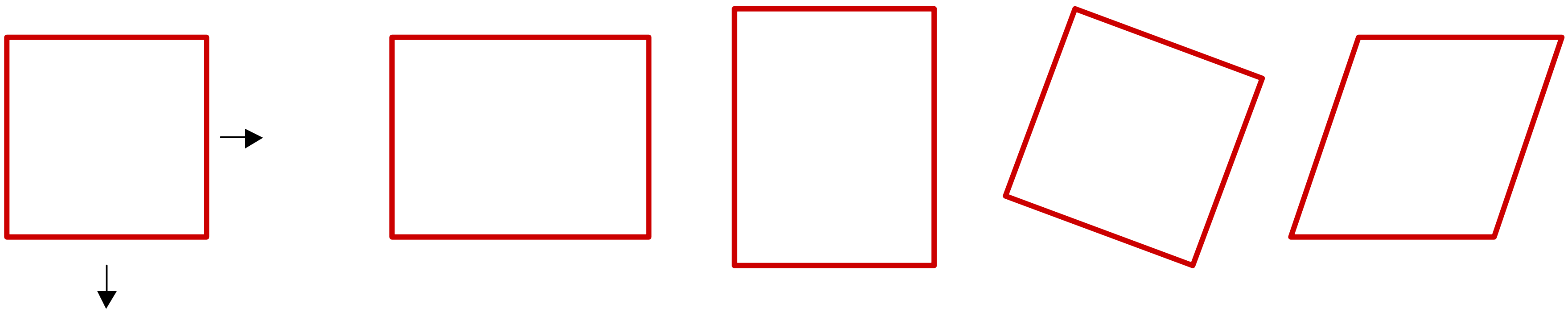
Generalization

- Transformations/warping of image

$$E(\mathbf{A}, \mathbf{h}) = \sum_{\mathbf{x} \in R} \left[I(\mathbf{A}\mathbf{x} + \mathbf{h}) - I_0(\mathbf{x}) \right]^2$$

Affine: $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \mathbf{h} = \begin{bmatrix} dx \\ dy \end{bmatrix}$

Generalization



Affine: $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\mathbf{h} = \begin{bmatrix} dx \\ dy \end{bmatrix}$

Example: Affine Motion

$$\boxed{\begin{aligned} u(x,y) &= a_1 + a_2 x + a_3 y \\ v(x,y) &= a_4 + a_5 x + a_6 y \end{aligned}}$$

Substituting into the B.C. Equation:

$$I_x \cdot u + I_y \cdot v + I_t \approx 0$$

$$\boxed{I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \approx 0}$$

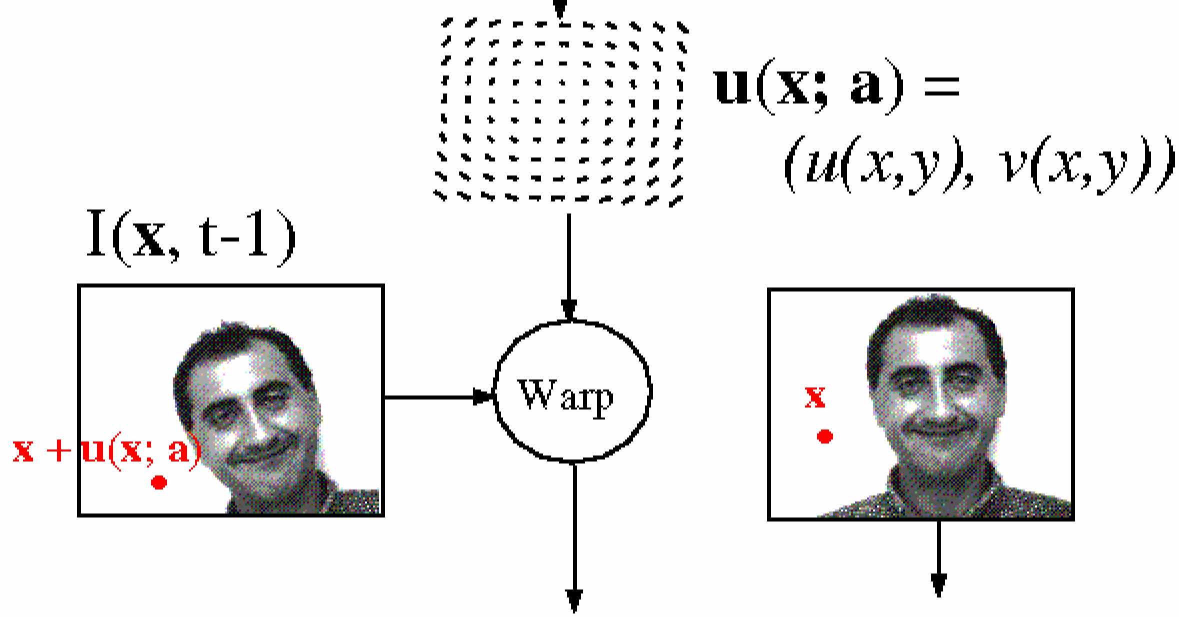
Each pixel provides 1 linear constraint in 6 *global* unknowns
(*minimum 6 pixels necessary*)

Least Square Minimization (over all pixels):

$$Err(\vec{a}) = \sum \left[I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \right]^2$$

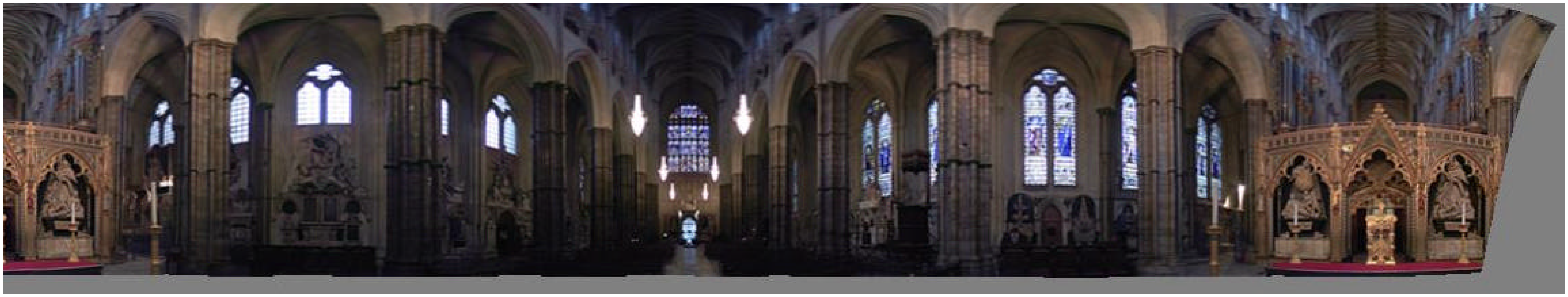
$$u(x, y) = a_0 + a_1 x + a_2 y$$

$$v(x, y) = a_3 + a_4 x + a_5 y$$



$$I(\mathbf{x}+\mathbf{u}(\mathbf{x}; \mathbf{a}), t-1) = I(\mathbf{x}, t)$$

(Brightness Constancy Assumption)



Generalization

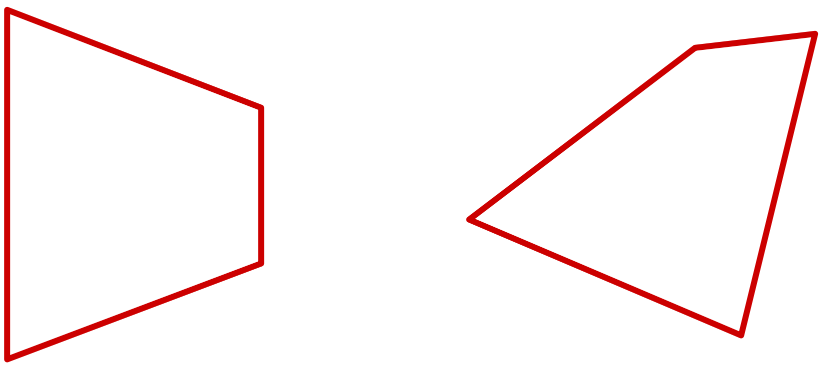
- Transformations/warping of image

$$E(\mathbf{A}) = \sum_{\mathbf{x} \in R} \left[I(\mathbf{A}\mathbf{x}) - I_0(\mathbf{x}) \right]^2$$

Planar perspective: $\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{bmatrix}$

Generalization

Affine +



$$\text{Planar perspective: } \mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{bmatrix}$$

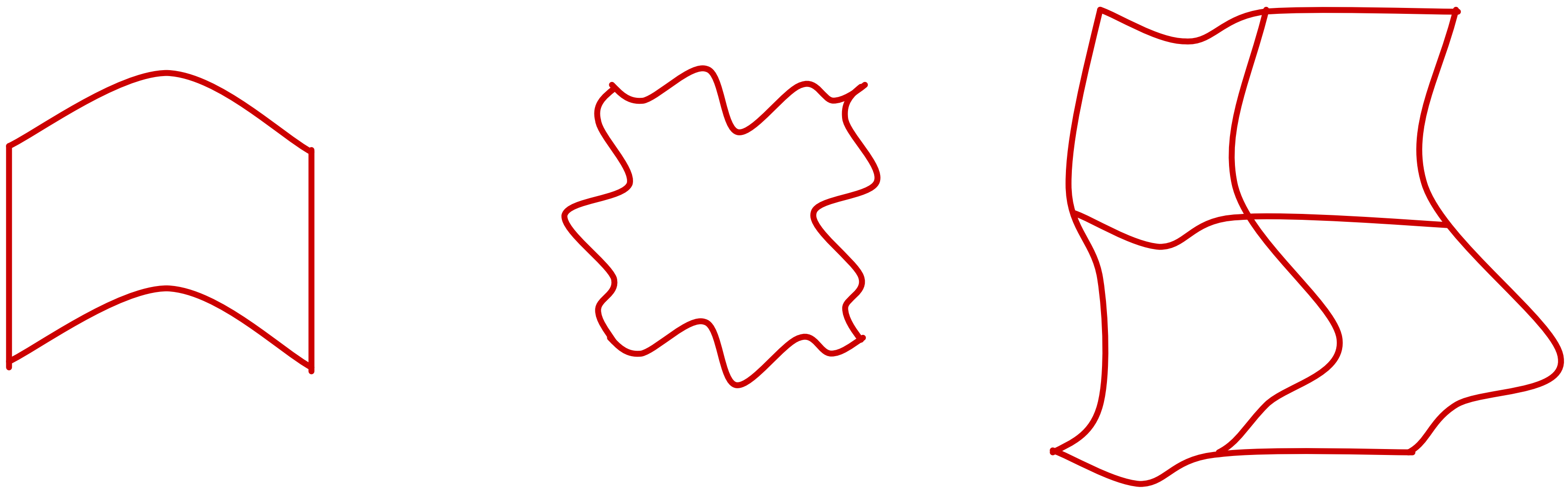
Generalization

- Transformations/warping of image

$$E(\mathbf{h}) = \sum_{\mathbf{x} \in R} \left[I(\mathbf{f}(\mathbf{x}, \mathbf{h})) - I_0(\mathbf{x}) \right]^2$$

Other parametrized transformations

Generalization



Other parametrized transformations

2D Motion Models summary

Quadratic – instantaneous approximation to planar motion

$$u = q_1 + q_2 x + q_3 y + q_7 x^2 + q_8 xy$$
$$v = q_4 + q_5 x + q_6 y + q_7 xy + q_8 y^2$$

Projective – exact planar motion

$$x' = \frac{h_1 + h_2 x + h_3 y}{h_7 + h_8 x + h_9 y}$$
$$y' = \frac{h_4 + h_5 x + h_6 y}{h_7 + h_8 x + h_9 y}$$

and

$$u = x' - x, \quad v = y' - y$$

3D Motion Models summary

Instantaneous camera motion:

Global parameters: $\Omega_X, \Omega_Y, \Omega_Z, T_X, T_Y, T_Z$

Local Parameter: $Z(x, y)$

$$u = -xy\Omega_X + (1 + x^2)\Omega_Y - y\Omega_Z + (T_X - T_Z x)/Z$$
$$v = -(1 + y^2)\Omega_X + xy\Omega_Y - x\Omega_Z + (T_Y - T_Z x)/Z$$

Homography+Epipole

Global parameters: $h_1, \ldots, h_9, t_1, t_2, t_3$

Local Parameter: $g(x, y)$

$$x' = \frac{h_1 x + h_2 y + h_3 + g\, t_1}{h_7 x + h_8 y + h_9 + g\, t_3}$$
$$y' = \frac{h_4 x + h_5 y + h_6 + g\, t_1}{h_7 x + h_8 y + h_9 + g\, t_3}$$
and : $\quad u = x' - x, \quad v = y' - y$

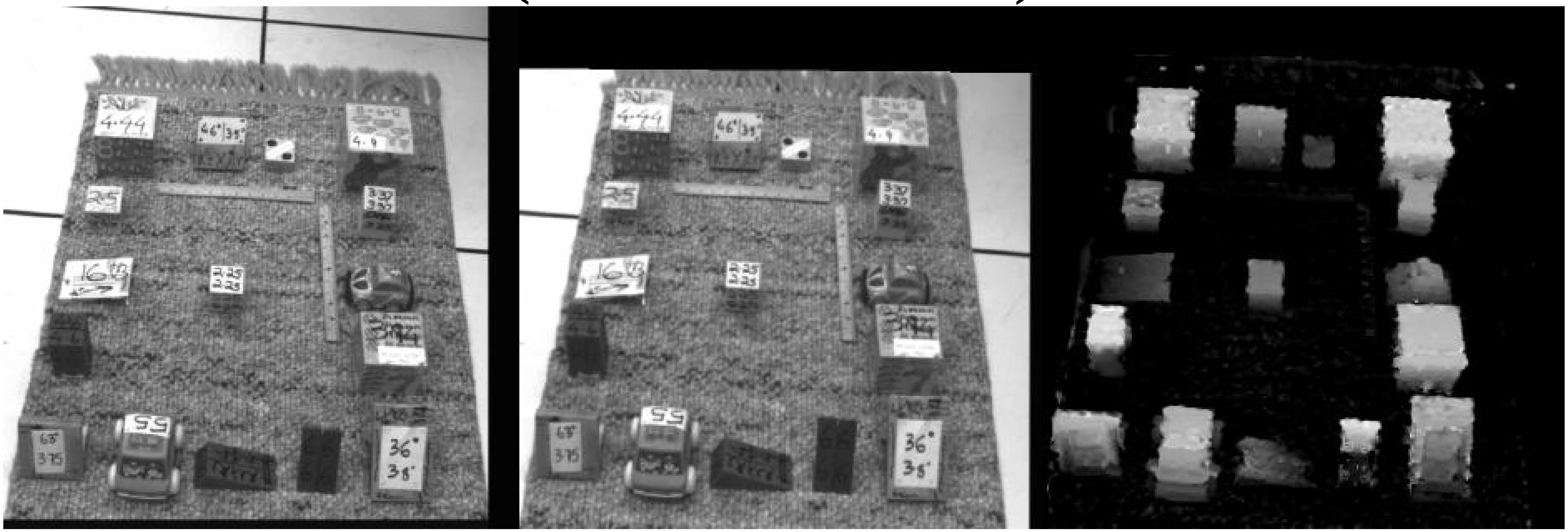
Residual Planar Parallax Motion

Global parameters: t_1, t_2, t_3

Local Parameter: $g(x, y)$

$$u = x^w - x = \frac{g}{1 + g\, t_3}(t_3 x - t_1)$$
$$v = y^w - x = \frac{g}{1 + g\, t_3}(t_3 y - t_2)$$

Residual Planar Parallax Motion (Plane+Parallax)

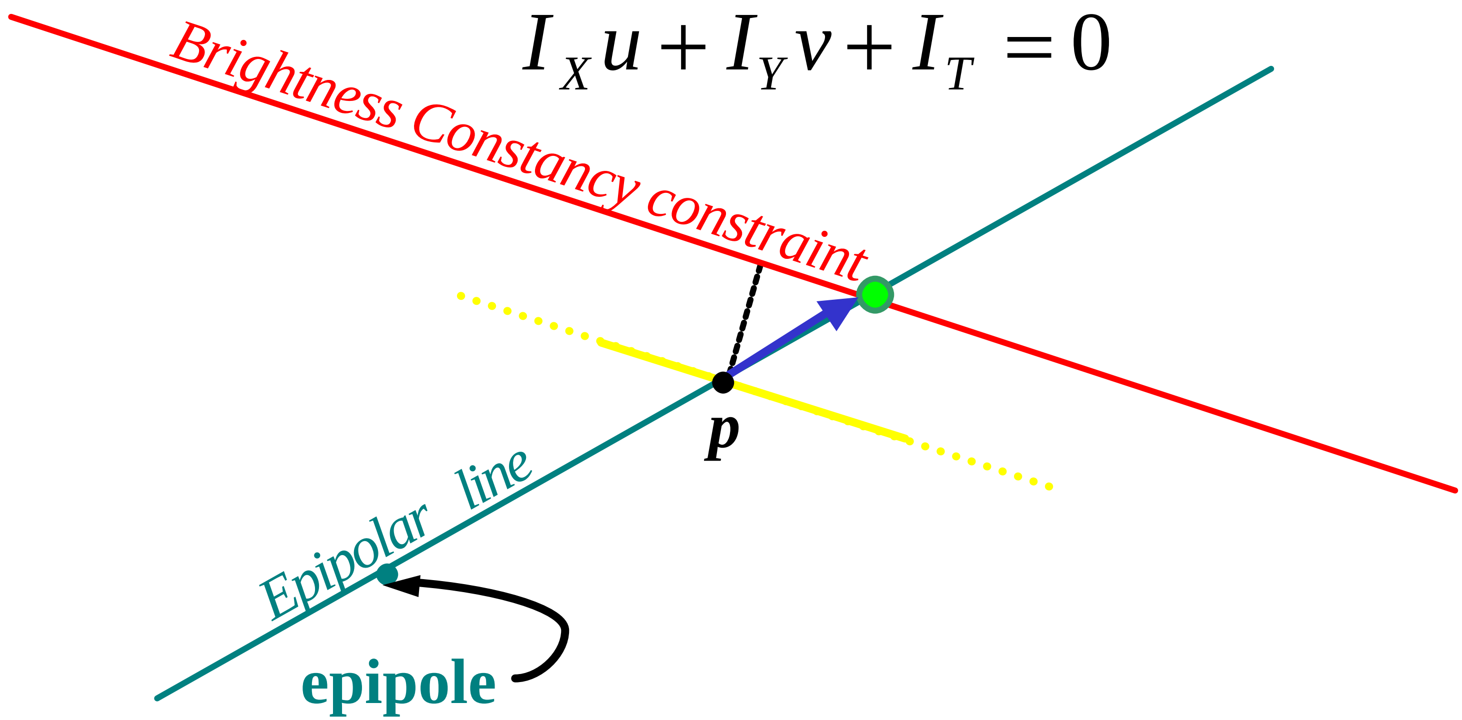


Original sequence **Plane-aligned sequence** **Recovered shape**

Block sequence from [Kumar-Anandan-Hanna'94]

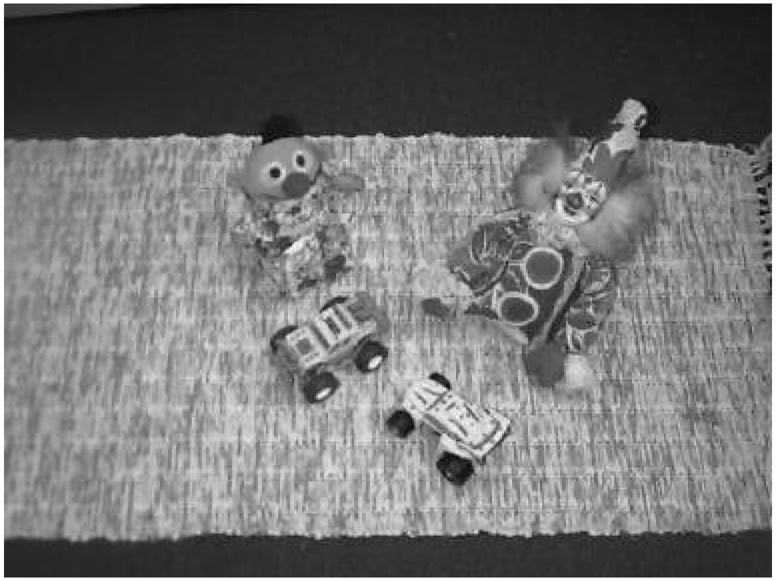
"Given two views where motion of points on a parametric surface has been compensated, the residual parallax is an epipolar field"

Residual Planar Parallax Motion

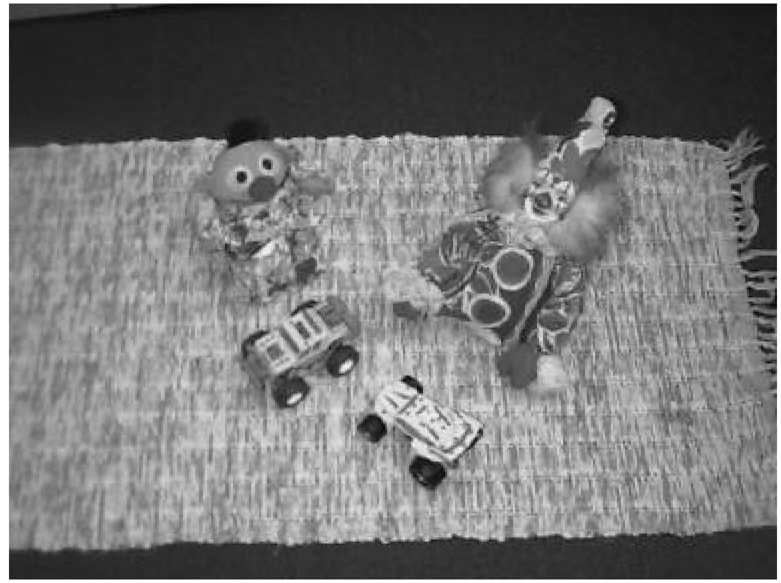


The intersection of the two line constraints <u>uniquely</u> defines the displacement.

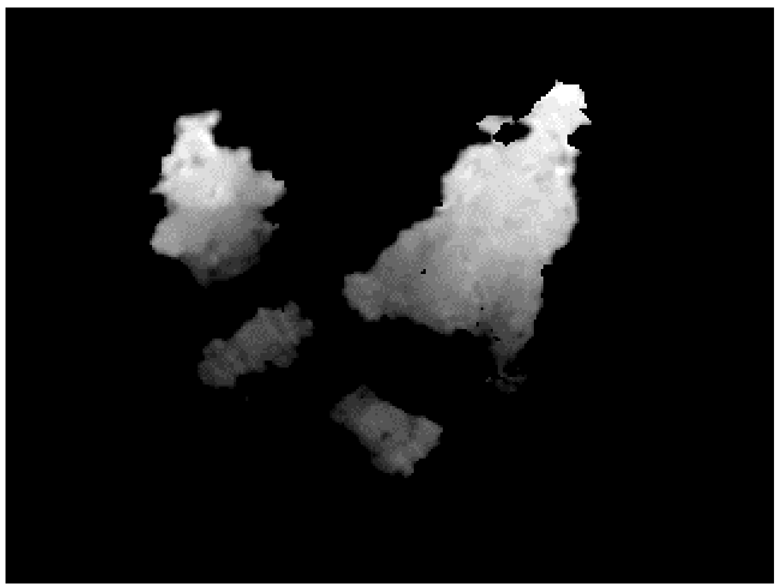
Dense 3D Reconstruction
(Plane+Parallax)



Original sequence



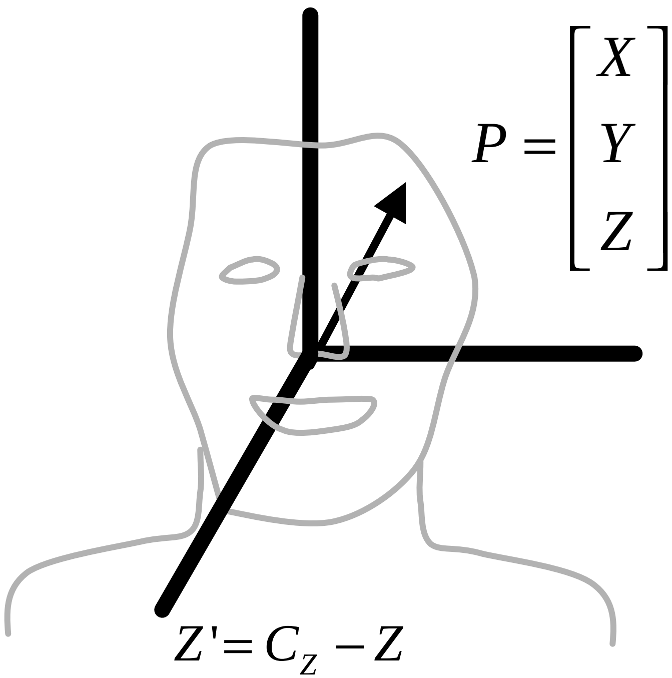
Plane-aligned sequence

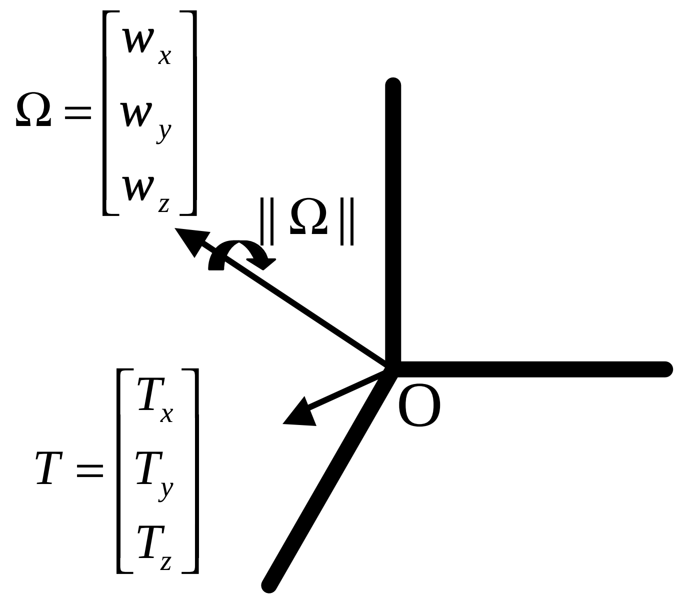


Recovered shape

Rigid pose estimation

- Head pose model: 6 DOF





Optic flow for rigid motion

- 3-D velocity:

$$V = T + \Omega \times P = T - \hat{\mathbf{P}}\Omega = \begin{bmatrix} \mathbf{I} & -\hat{\mathbf{P}} \end{bmatrix} \begin{bmatrix} T \\ \Omega \end{bmatrix}$$

$$\hat{\mathbf{P}} = [\mathbf{P}_x]$$
(skew-sym.)

$$V = \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -Y \\ 0 & 1 & 0 & -Z & 0 & X \\ 0 & 0 & 1 & Y & -X & 0 \end{bmatrix} \begin{bmatrix} T \\ \Omega \end{bmatrix}$$

(See Horn, 17.2)

Optic flow for rigid motion

- Perspective projection

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

Optic flow for rigid motion

- Combine

$$V = \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -Y \\ 0 & 1 & 0 & -Z & 0 & X \\ 0 & 0 & 1 & Y & -X & 0 \end{bmatrix} \begin{bmatrix} T \\ \Omega \end{bmatrix}$$

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}$$

Optic flow for rigid motion

- Rigid Motion (for small v): $\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \mathbf{H} \begin{bmatrix} T \\ \Omega \end{bmatrix}$

$$\mathbf{H} = \begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -Y \\ 0 & 1 & 0 & -Z & 0 & X \\ 0 & 0 & 1 & Y & -X & 0 \end{bmatrix}$$

Perspective projection of 3-D velocity

3-D velocity at point P

Hard to solve with just optic flow vectors! (but see Horn 17.3-17.5).

Instead, ***Direct method*** combines constraint linearly with BCCE!

Direct Rigid Motion Estimation

- Brightness Change Constraint

$$I(x, y, t) = I(x + v_x, y + v_y, t + 1)$$

$$\frac{dI}{dx} v_x + \frac{dI}{dy} v_y + \frac{dI}{dt} = 0$$

$$\left[-\frac{dI}{dt} \right] = \left[\frac{dI}{dx} \quad \frac{dI}{dy} \right] \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

Direct Rigid Motion Estimation

- Brightness Change Constraint

$$\begin{bmatrix} -\frac{dI}{dt} \end{bmatrix} = \begin{bmatrix} \frac{dI}{dx} & \frac{dI}{dy} \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

- Rigid Motion Model

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \mathbf{H} \begin{bmatrix} T \\ \Omega \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -Y \\ 0 & 1 & 0 & -Z & 0 & X \\ 0 & 0 & 1 & Y & -X & 0 \end{bmatrix}$$

Direct Motion Estimation

- One equation per pixel:

$$\begin{bmatrix} -\frac{dI}{dt} \end{bmatrix} = \begin{bmatrix} \frac{dI}{dx} & \frac{dI}{dy} \end{bmatrix} \begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -Y \\ 0 & 1 & 0 & -Z & 0 & X \\ 0 & 0 & 1 & Y & -X & 0 \end{bmatrix} \begin{bmatrix} \mathrm{T} \\ \Omega \end{bmatrix}$$

First, convert X,Y from screen coordinates to pixel coordinates….

Direct Motion Estimation

- One equation per pixel:

$$\begin{bmatrix} -\dfrac{dI}{dt} \end{bmatrix} = \begin{bmatrix} \dfrac{dI}{dx} & \dfrac{dI}{dy} \end{bmatrix} \begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -yZ'/f \\ 0 & 1 & 0 & -Z & 0 & xZ'/f \\ 0 & 0 & 1 & yZ'/f & -xZ'/f & 0 \end{bmatrix} \begin{bmatrix} \mathrm{T} \\ \Omega \end{bmatrix}$$

Direct Motion Estimation

- One equation per pixel:

$$\left[-\frac{dI}{dt}\right]=\begin{bmatrix}\frac{dI}{dx} & \frac{dI}{dy}\end{bmatrix}\begin{bmatrix} f & 0 & -x \\ 0 & f & -y \end{bmatrix}\underbrace{\frac{1}{Z'}}\begin{bmatrix} 1 & 0 & 0 & 0 & Z & -yZ'/f \\ 0 & 1 & 0 & -Z & 0 & xZ'/f \\ 0 & 0 & 1 & \underbrace{yZ'/f} & -xZ'/f & 0 \end{bmatrix}\begin{bmatrix} \mathrm{T} \\ \Omega \end{bmatrix}$$

- Still hard!
- Z unknown; assume surface shape…
  - Negahdaripour & Horn - Planar
  - Black and Yacoob - Affine
  - Basu and Pentland; Bregler and Malik - Ellipsoidal
  - Essa et al. - Polygonal approximation
  - ...

“Direct Depth”

Use real-time stereo!

- Gives Z directly; no approximate model needed
- Express Direct Constraint on Depth Gradient

$$I(x,y,t) = I(x+v_x, y+v_y, t+1)$$

$$Z(x,y,t) = Z(x+v_x, y+v_y, t+1) - v_z$$

$$\frac{dZ}{dx}v_x + \frac{dZ}{dy}v_y + \frac{dZ}{dt} - v_z = 0$$

Direct Depth

3-D Depth and Brightness Constraint Equations:

- Orthographic

$$\begin{bmatrix} -dI/dt \\ -dZ/dt \end{bmatrix} = \begin{bmatrix} dI/dx & dI/dy & 0 \\ dZ/dx & dZ/dy & -1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$

- Perspective

$$\begin{bmatrix} -dI/dt \\ -dZ/dt \end{bmatrix} = \begin{bmatrix} fdI/dx & fdI/dy & -ydI/dy - xdI/dx \\ fdZ/dx & fdZ/dy & -1 \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$$

Direct Depth

Combined with rigid motion model:

- Orthographic

$$\begin{bmatrix} -dI/dt \\ -dZ/dt \end{bmatrix} = \begin{bmatrix} dI/dx & dI/dy & 0 \\ dZ/dx & dZ/dy & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -y \\ 0 & 1 & 0 & -Z & 0 & x \\ 0 & 0 & 1 & y & -x & 0 \end{bmatrix} \begin{bmatrix} \mathrm{T} \\ \Omega \end{bmatrix}$$

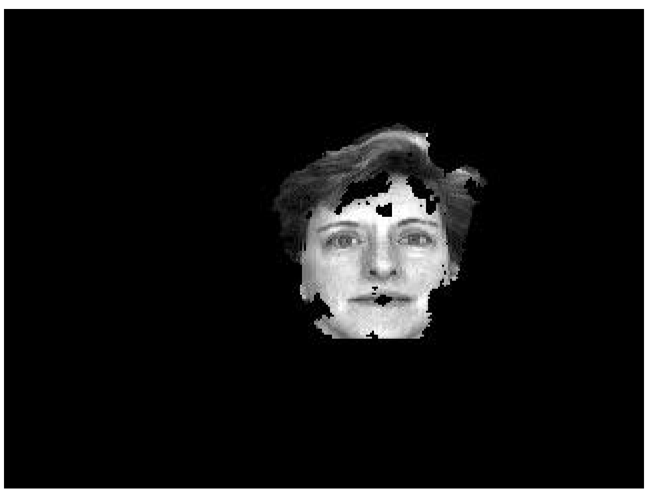
- Perspective

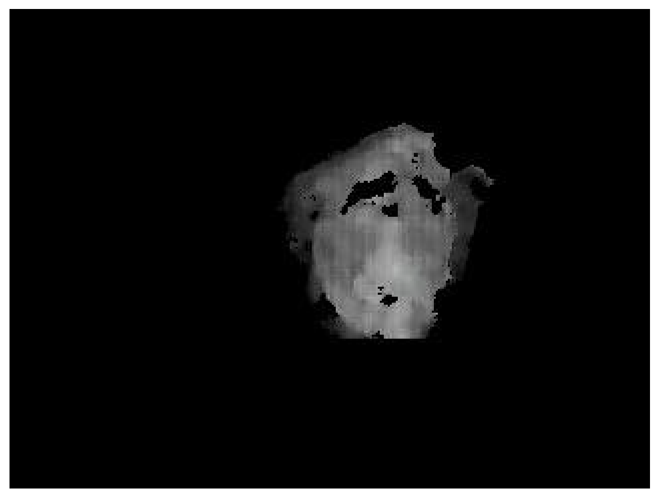
$$\begin{bmatrix} -dI/dt \\ -dZ/dt \end{bmatrix} = \begin{bmatrix} fdI/dx & fdI/dy & -ydI/dy - xdI/dx \\ fdZ/dx & fdZ/dy & -1 \end{bmatrix} \frac{1}{Z'} \begin{bmatrix} 1 & 0 & 0 & 0 & Z & -yZ'/f \\ 0 & 1 & 0 & -Z & 0 & xZ'/f \\ 0 & 0 & 1 & yZ'/f & -xZ'/f & 0 \end{bmatrix} \begin{bmatrix} \mathrm{T} \\ \Omega \end{bmatrix}$$

One system per pixel, same T,Ω. Solve with QR or SVD.

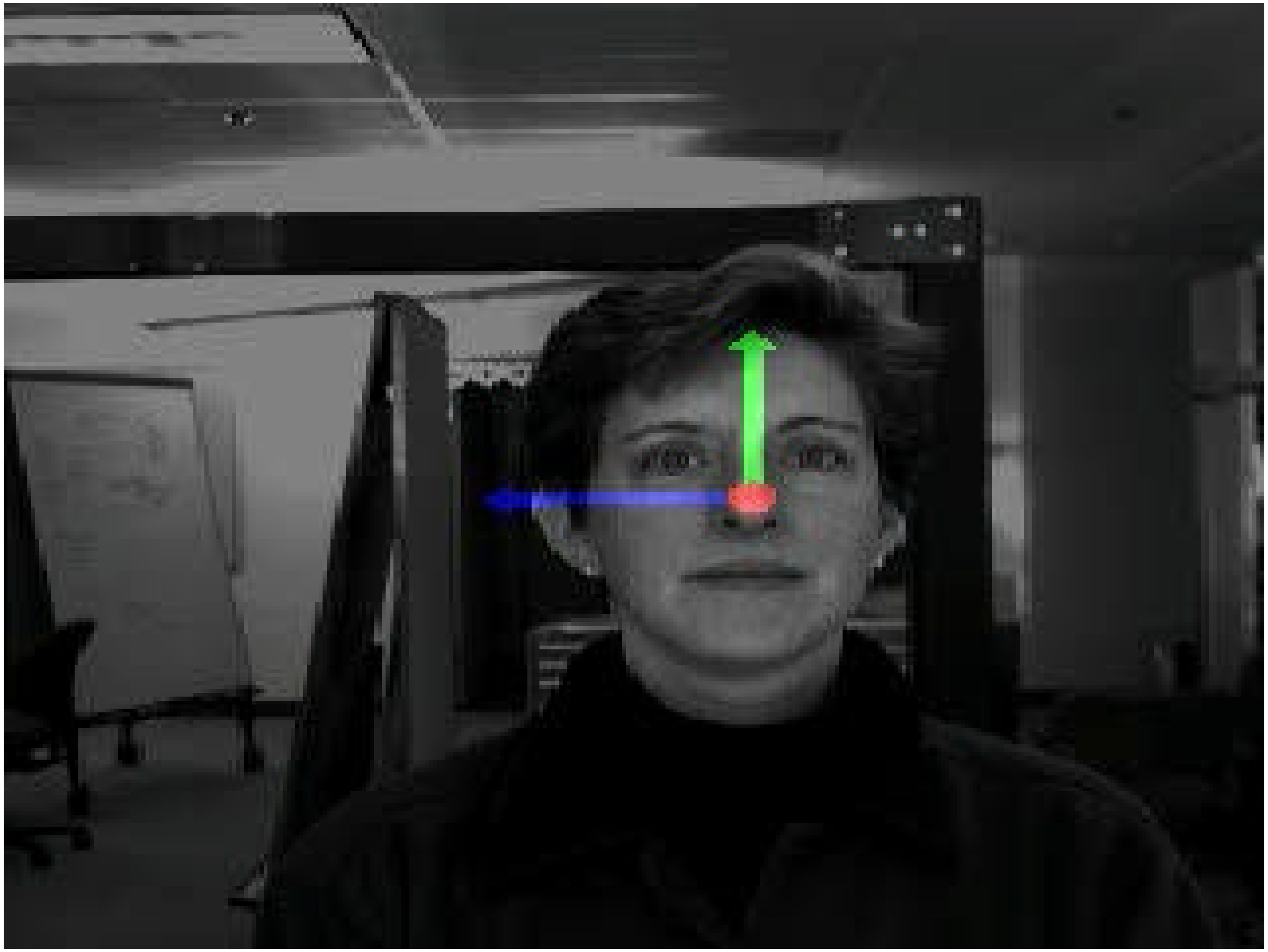
[Harville et. al]

Input to Pose Tracking



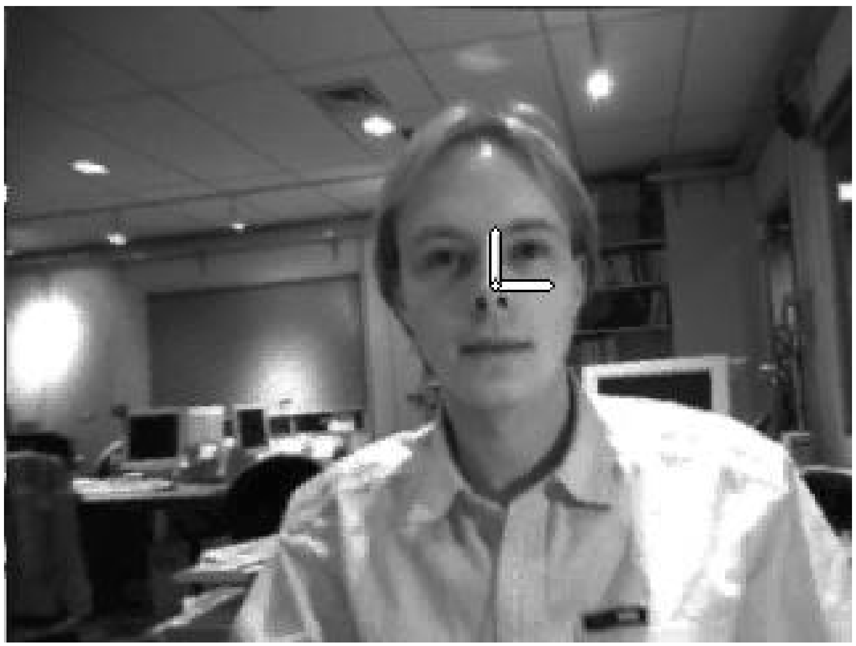


Pose Results



Application

Track users’ head gaze for hands-free pointing…



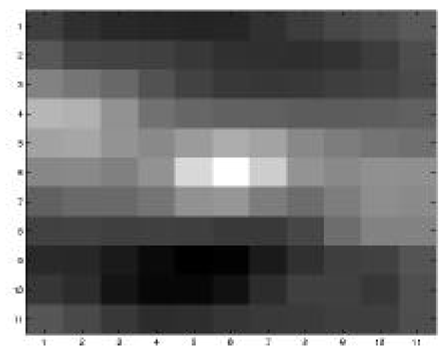
Topics

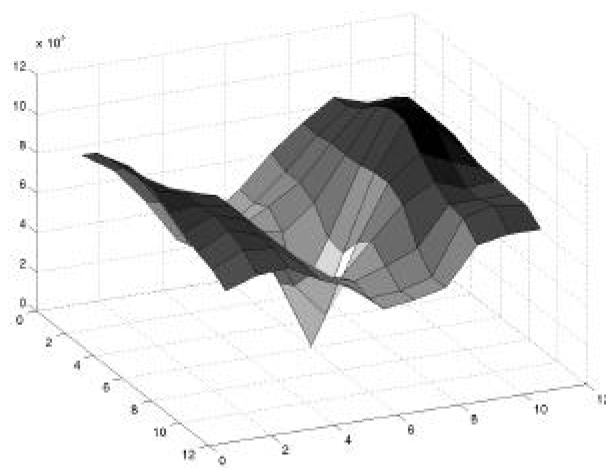
- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- Direct depth
- SSD tracking
- Robust flow
- Bayesian flow

Correlation and SSD

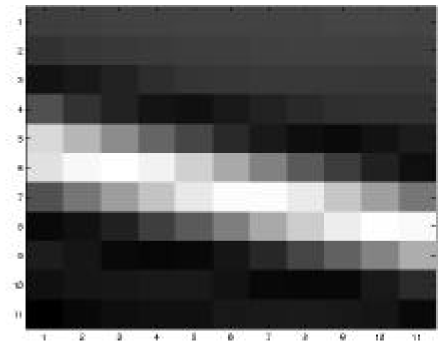
- For large displacements, do template matching as was used in stereo disparity search.
  - Define a small area around a pixel as the template
  - Match the template against each pixel within a search area in next image.
  - Use a match measure such as correlation, normalized correlation, or sum-of-squares difference
  - Choose the maximum (or minimum) as the match
  - Sub-pixel interpolation also possible

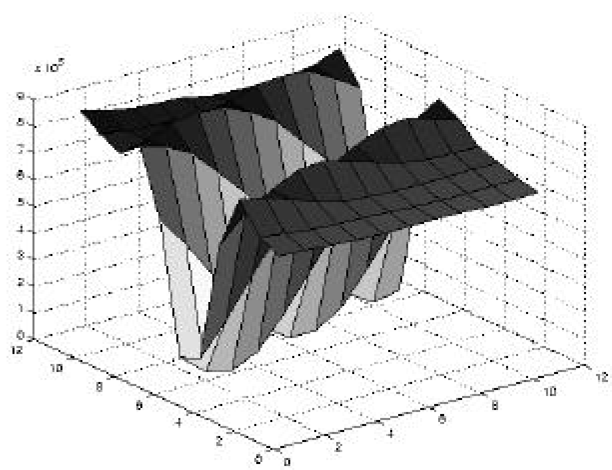
SSD Surface – Textured area



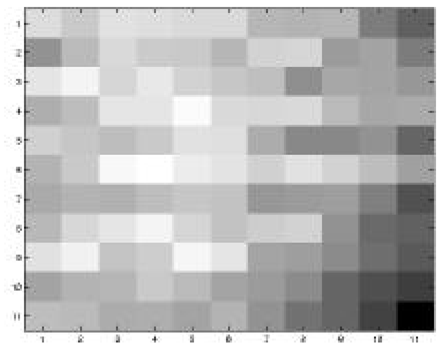


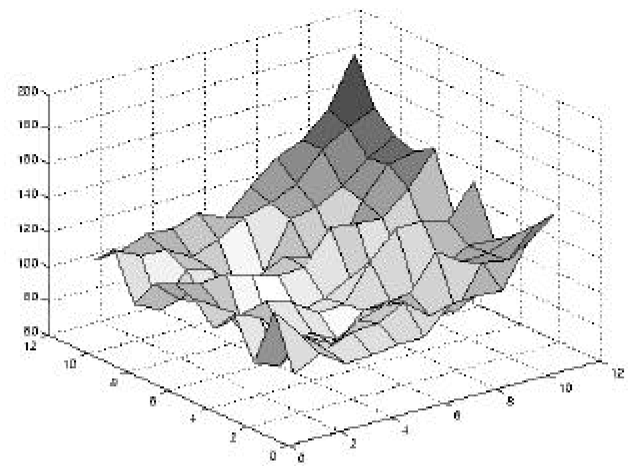
SSD Surface -- Edge





SSD Surface – homogeneous area





Discrete Search vs. Gradient Based Estimation

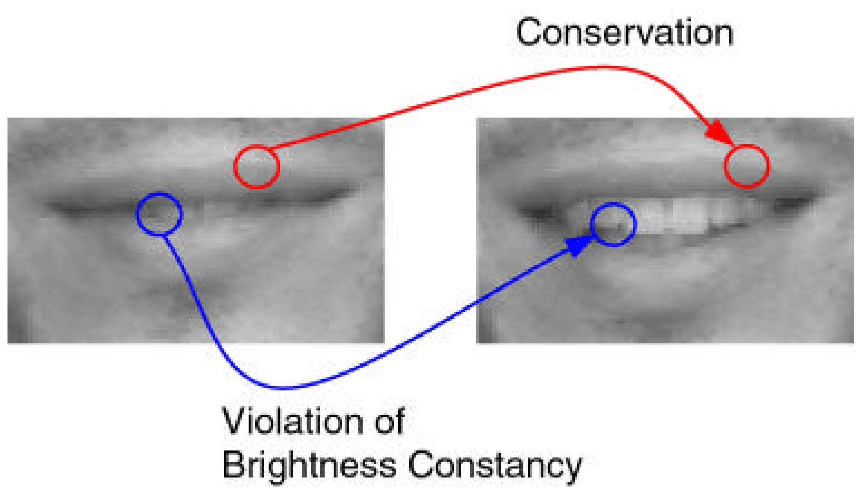
Consider image I translated by u_0, v_0

$$\boxed{\begin{aligned} I_0(x,y) &= I(x,y) \\ I_1(x+u_0, y+v_0) &= I(x,y) + h_1(x,y) \end{aligned}}$$

$$\boxed{\begin{aligned} E(u,v) &= \sum_{x,y} (I(x,y) - I_1(x+u, y+v))^2 \\ &= \sum_{x,y} (I(x,y) - I(x-u_0+u, y-v_0+v) - h_1(x,y))^2 \end{aligned}}$$

Discrete search simply searches for the best estimate.
Gradient method linearizes the intensity function and solves for the estimate

Problem



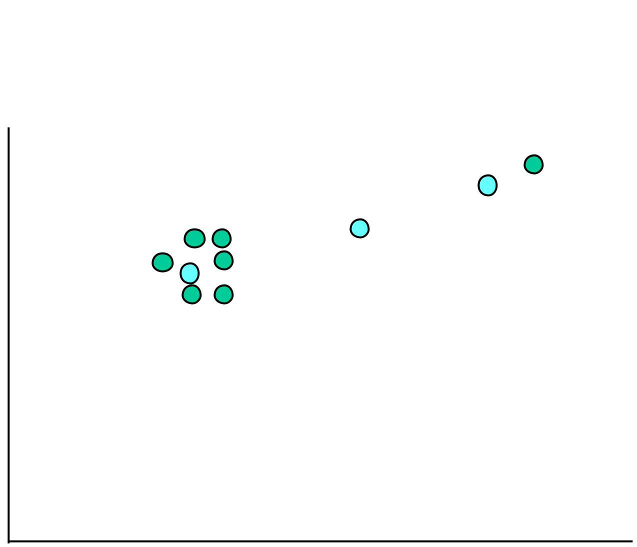
Violations of brightness constancy result in large residual errors:

$$\left| \mathrm{I}_x u(\mathbf{x};\mathbf{a}) + \mathrm{I}_y v(\mathbf{x};\mathbf{a}) + \mathrm{I}_t \right|_\theta$$

Choose θ to be insensitive to *outliers*

Robust Estimation

Standard Least Squares Estimation allows too much influence for outlying points



$$E(m) = \sum_i r\,(x_i)$$

$$r\,(x_i) = (x_i - m)^2$$

Influence $y\,(x) = \dfrac{\partial r}{\partial x} = (x_i - m)$

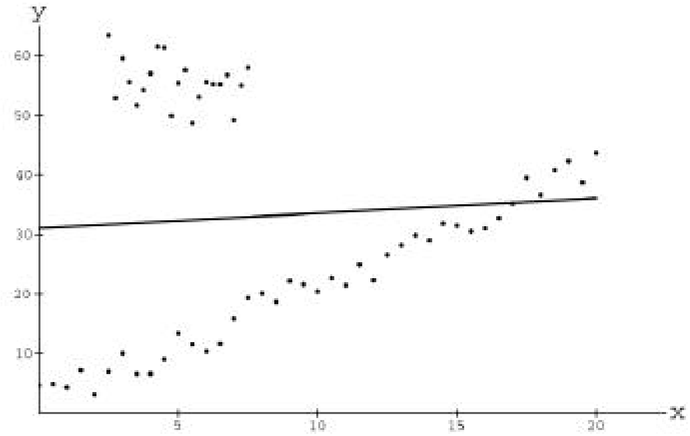
Robust Estimation

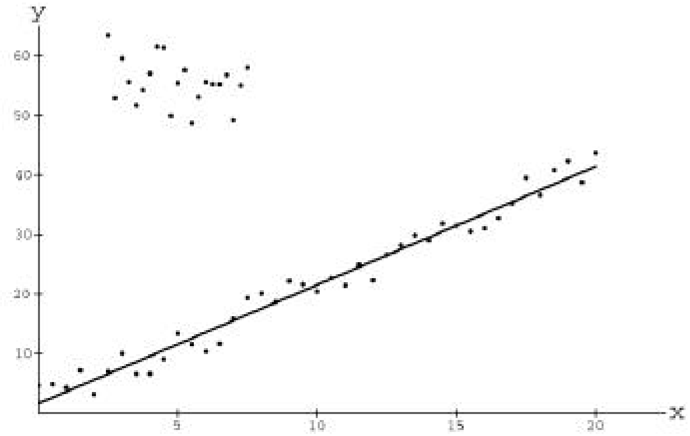
Goal:

- Recover the best fit to the majority of the data
- detect and reject outliers

Linear regression:

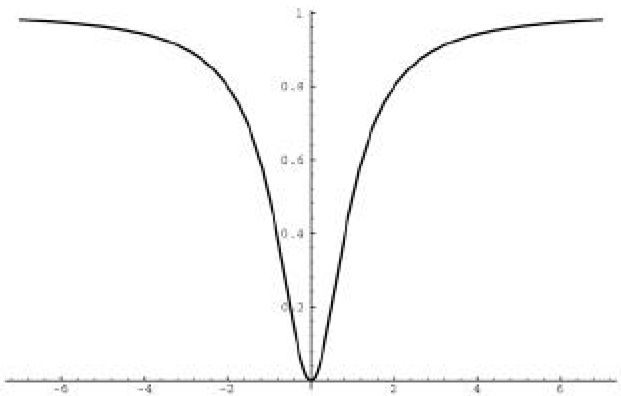
$$E(a_1, a_2) = \sum_i r((a_1 x_i + a_2) - d_i, s)$$

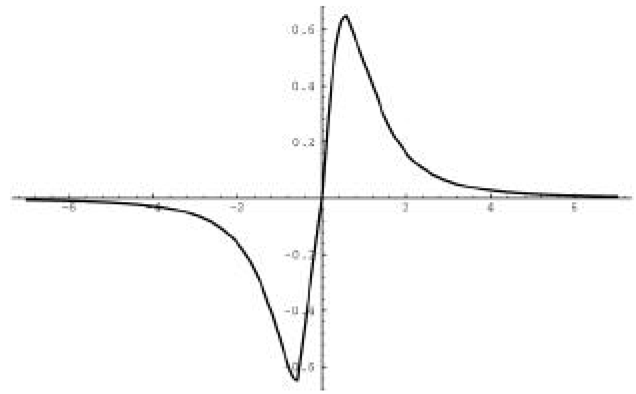




Robust Estimation

Quadratic θ function gives too much weight to outliers.





$$r(r, s) = \frac{r^2}{s^2 + r^2} \qquad y(r, s) = \frac{2rs^2}{(s^2 + r^2)^2}$$

Robust Estimation

$$E_d(u_s, v_s) = \sum r\left(I_x u_s + I_y v_s + I_t\right)$$ Robust gradient constraint

$$E_d(u_s, v_s) = \sum r\left(I(x, y) - J(x + u_s, y + v_s)\right)$$ Robust SSD

Minimize via coordinate descent.

- Non-convex
- Use a continuation method with deterministic annealing.
- Coarse-to-fine
- IRLS

Iteratively Reweighted Least-Squares

Robust Minimization (over all pixels):

$$Err(\vec{a}) = \sum W(x,y)\left[I_x(x,y)u(x,y) + I_y(x,y)v(x,y) + I_t(x,y)\right]^2$$

Reweight each pixel based on previous residual.

An Outlier Measure is thus:

$O(x,y)$ inversely related to $(1/W(x,y))$

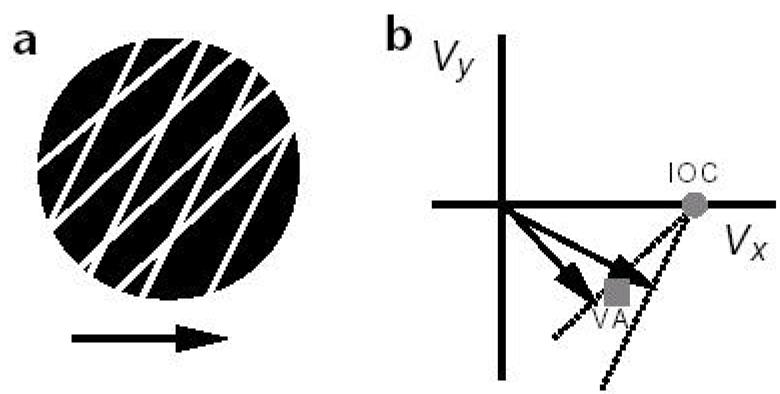
$$O(x,y) \propto \frac{I_t\left|\nabla I\right|}{\left|\nabla I\right|^2 + e}$$

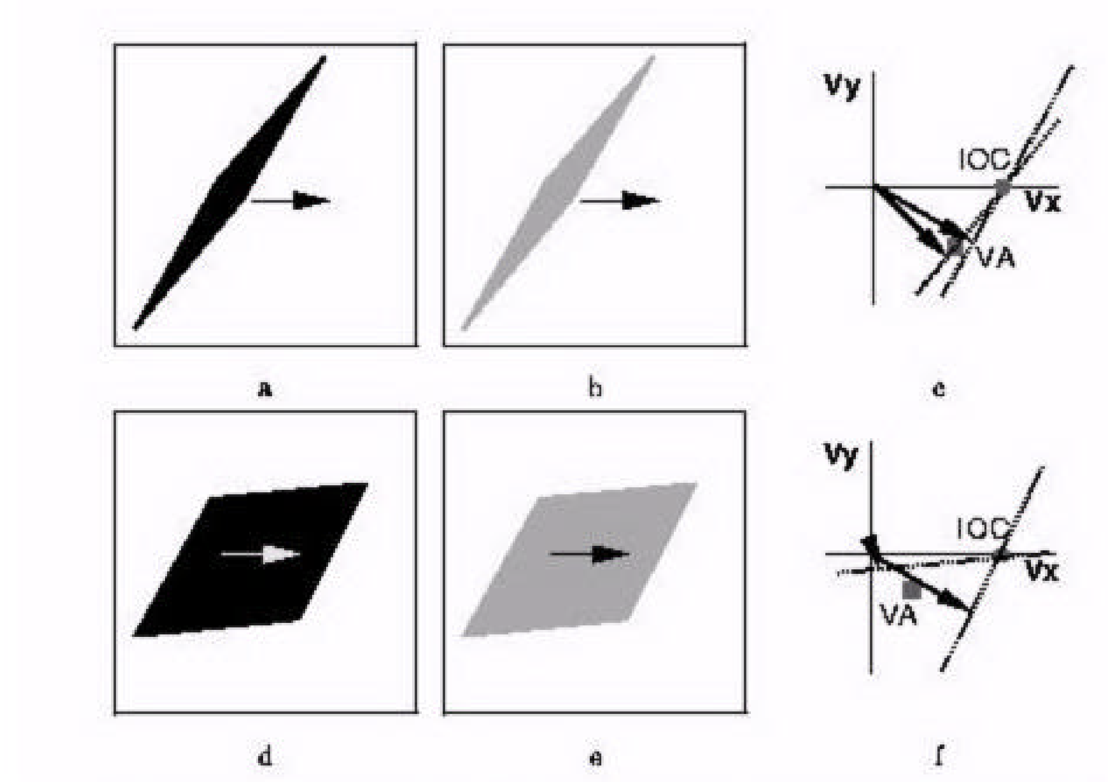
Topics

- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- Direct depth
- SSD tracking
- Robust flow
- Bayesian flow

Bayesian Optic Flow

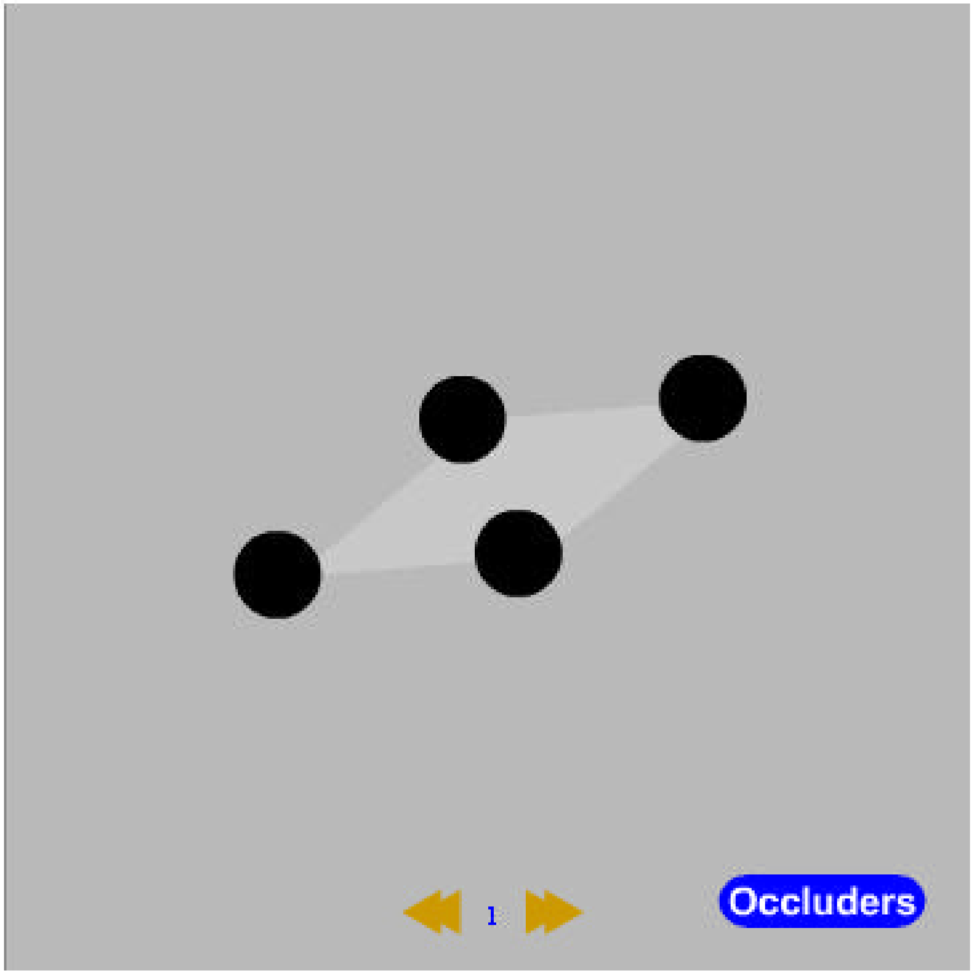
- Some low-level human motion illusions can be explained by adding an uncertianty model to Lucas-Kanade tracking
- Theories from Psychology about normal flow fusion:
  - (VA) vector average (of normal motions)
  - (IOC) intersection of constraints (e.g., Lucas-Kanade):

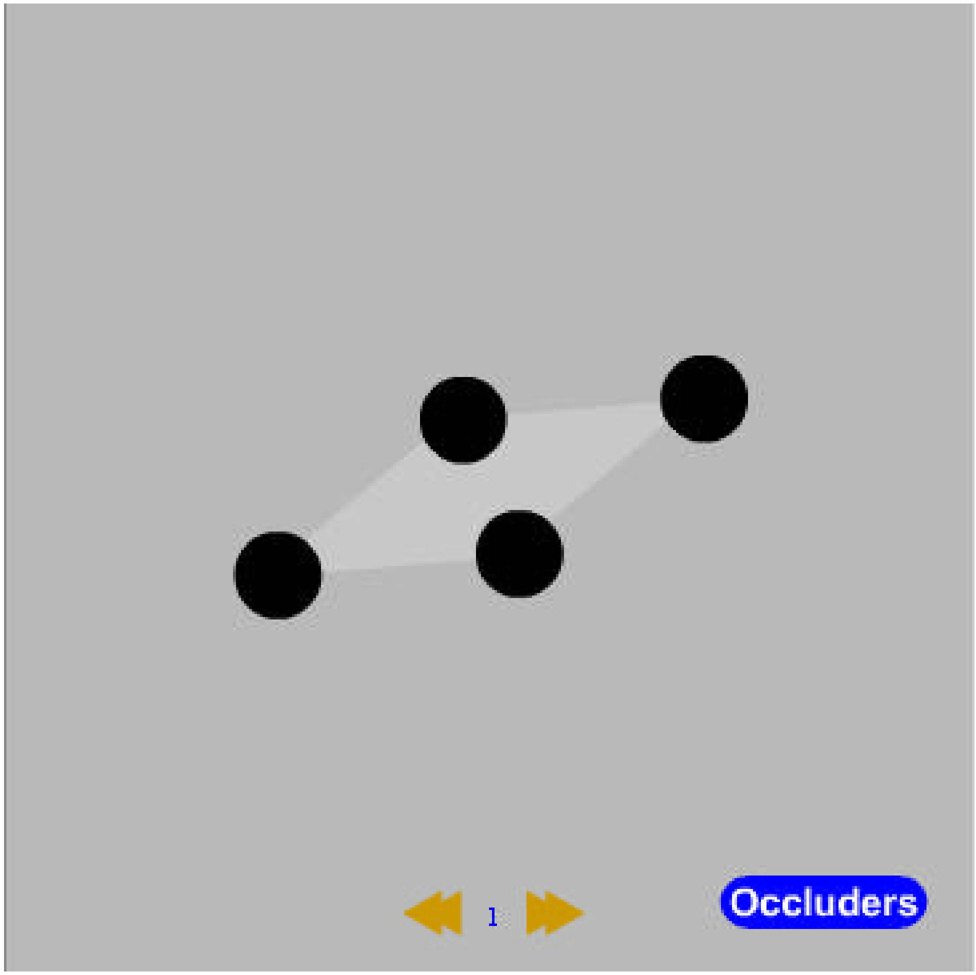


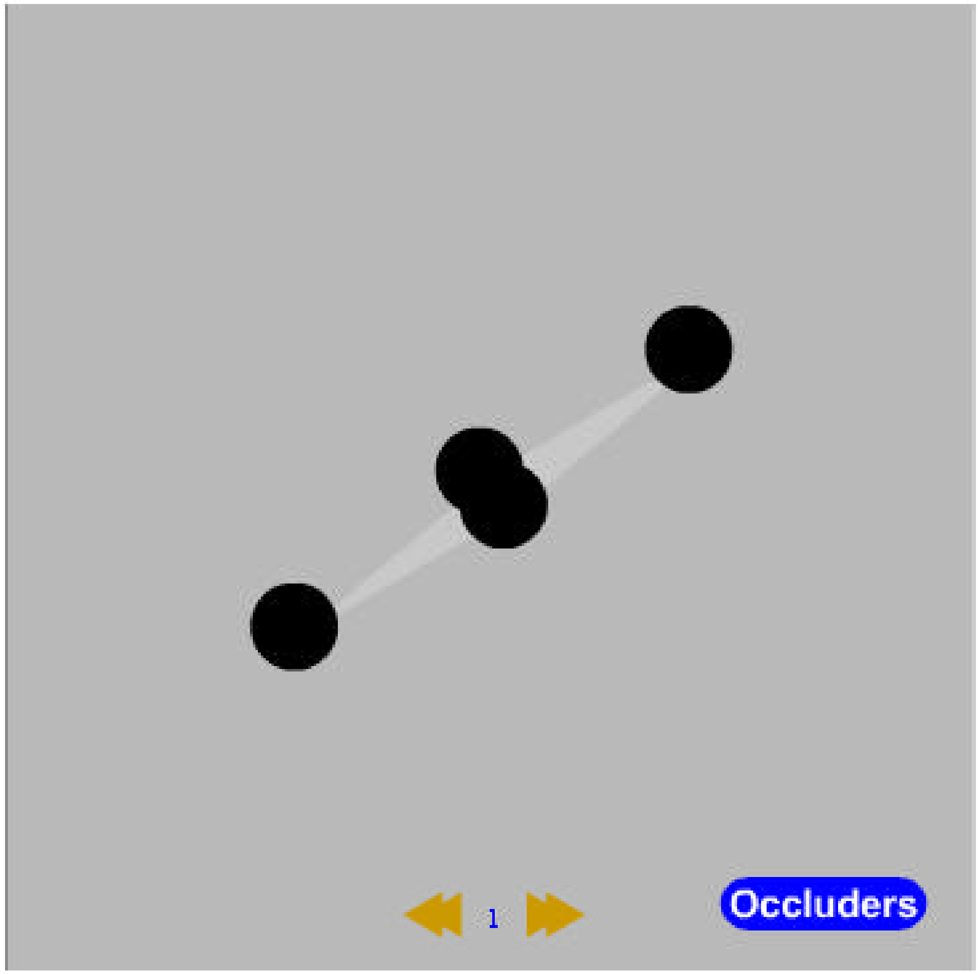


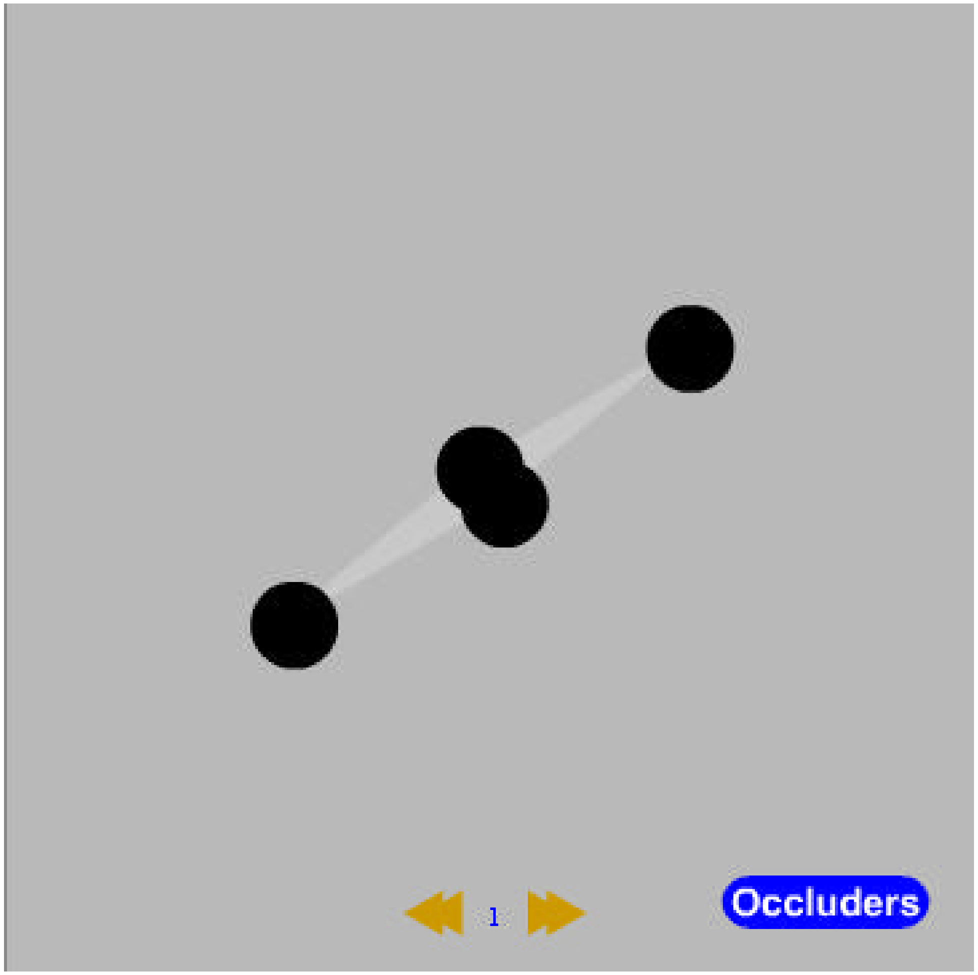
This demo illustrates the insufficiency of either VA, IOC or feature tracking to explain human perception of a horizontially moving rhombus. *a.* A ``narrow'' rhombus at high contrast appears to move horizontally (consistent with IOC or feature tracking). *b.* A ``narrow'' rhombus at low contrast appears to move diagonally (consistent with VA). *c.* Velocity space constraints for a narrow rhombus. *d-e.* A ``fat'' rhombus at low and high contrast appears to move horizontally (consistent with IOC or feature tracking). *f.* Velocity space constraints of a fat rhombus.

http://www.cs.huji.ac.il/~yweiss/Rhombus.


1
Occluders


1
Occluders


1
Occluders


1
Occluders

Brightness constancy with noise:

$$I(x,y,t) = I(x + v_x\Delta t,\ y + v_y\Delta t,\ t + \Delta t) + \eta$$

Assume Gaussian noise, smooth surfaces, locally constant; take first order linear approximation:

$$P(I(x_i,y_i,t)|v_i) \propto$$
$$\exp\left(-\frac{1}{2\sigma^2} \int_{x,y} w_i(x,y)\ (I_x(x,y,t)v_x + I_y(x,y,t)v_y + I_t(x,y,t))^2\ \mathrm{d}x\,\mathrm{d}y \right)$$

Prior favoring slow speeds:

$$P(v) \propto \exp(-\|v\|^2/2\sigma_p{}^2).$$

Assume noise is independent across location; apply Bayes:

$$P(v|I) \propto P(v) \prod_i P(I(x_i,y_i,t)\ |v),$$

With constant window w=1,

$$P(v|I) \propto \exp\left(-\|v\|^2/2\sigma_p{}^2 - \frac{1}{2\sigma^2} \int_{x,y} (I(x,y)\ v_x + I_y(x,y)v_y + I_t)^2\ \mathrm{d}x\,\mathrm{d}y \right)$$

Form ‘normal equations’ to arrive at….

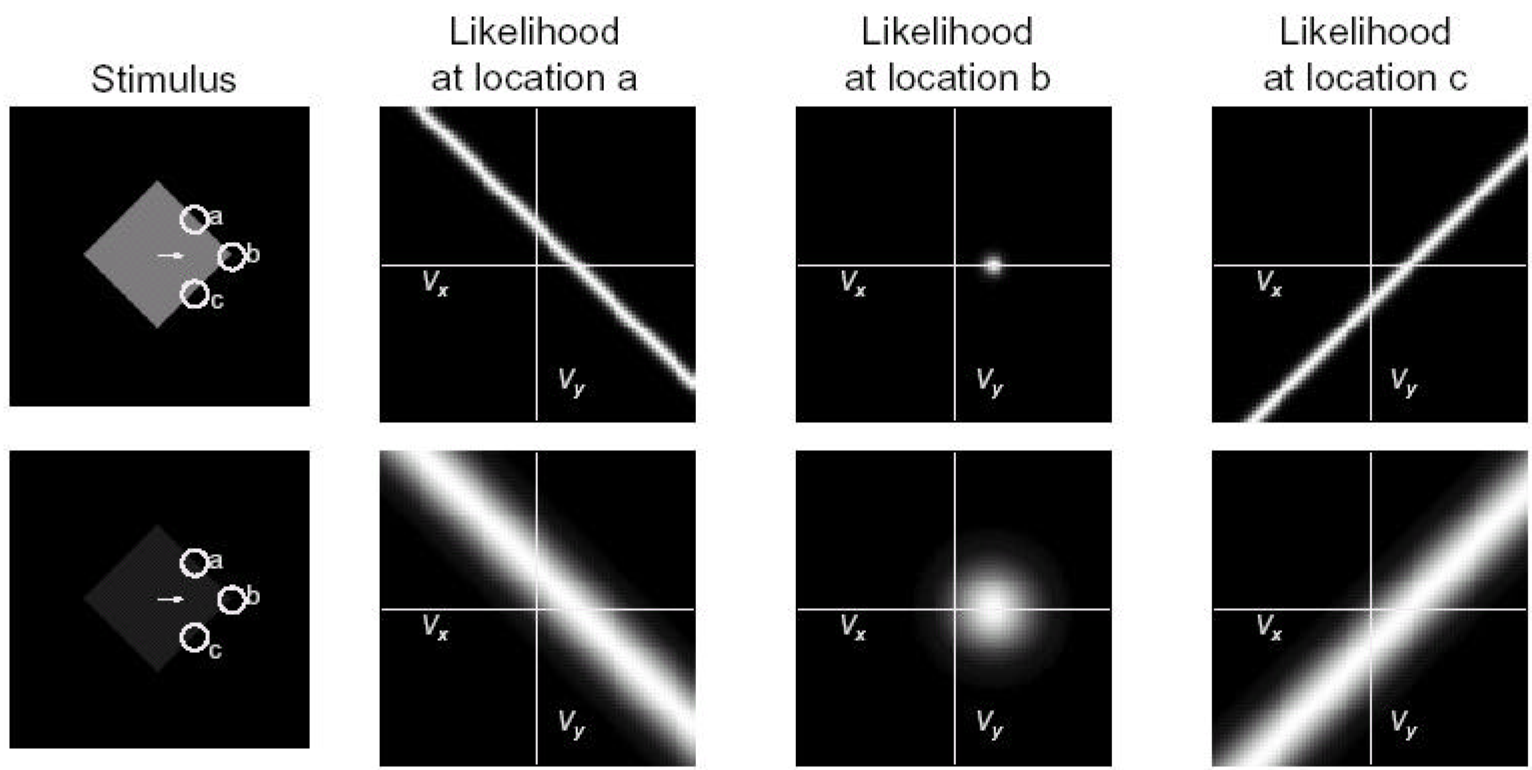
Lucas-Kanade with uncertainty:

$$v^* = -\begin{pmatrix} \sum I_x{}^2 + \dfrac{\sigma^2}{\sigma_p{}^2} & \sum I_x I_y \\ \sum I_x I_y & \sum I_y{}^2 + \dfrac{\sigma^2}{\sigma_p{}^2} \end{pmatrix}^{-1} \begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}$$

One parameter: ratio of observation and prior gaussian spread.

http://www.cs.huji.ac.il/~yweiss/Rhombus

[Weiss, Simoncelli, Adelson Nature Neuroscience 2002]


Stimulus
Likelihood at location a
Likelihood at location b
Likelihood at location c
a
b
c
V_x
V_y

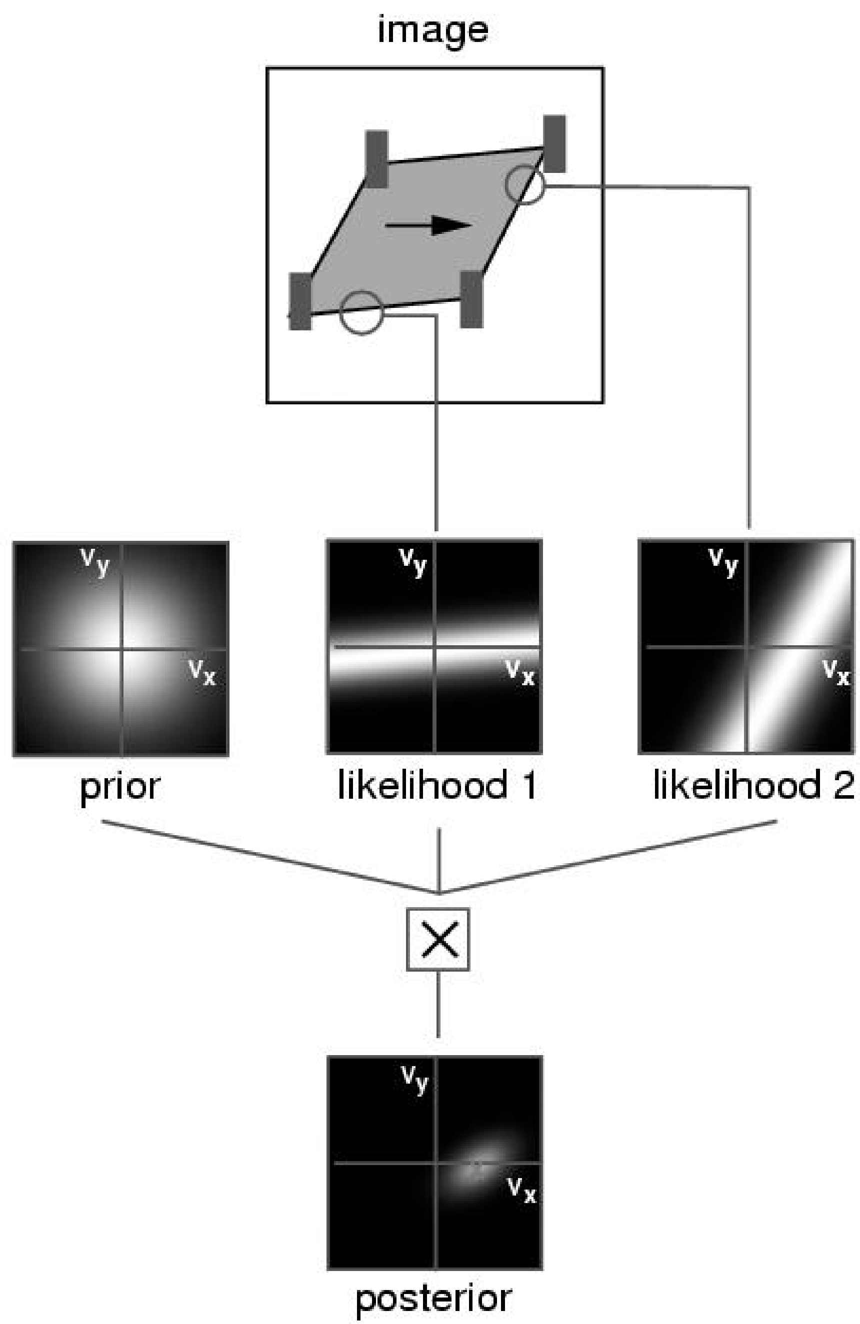


Figure 4: The response of the Bayesian estimator to a fat rhombus. (replotted from Weiss and Adelson 98)

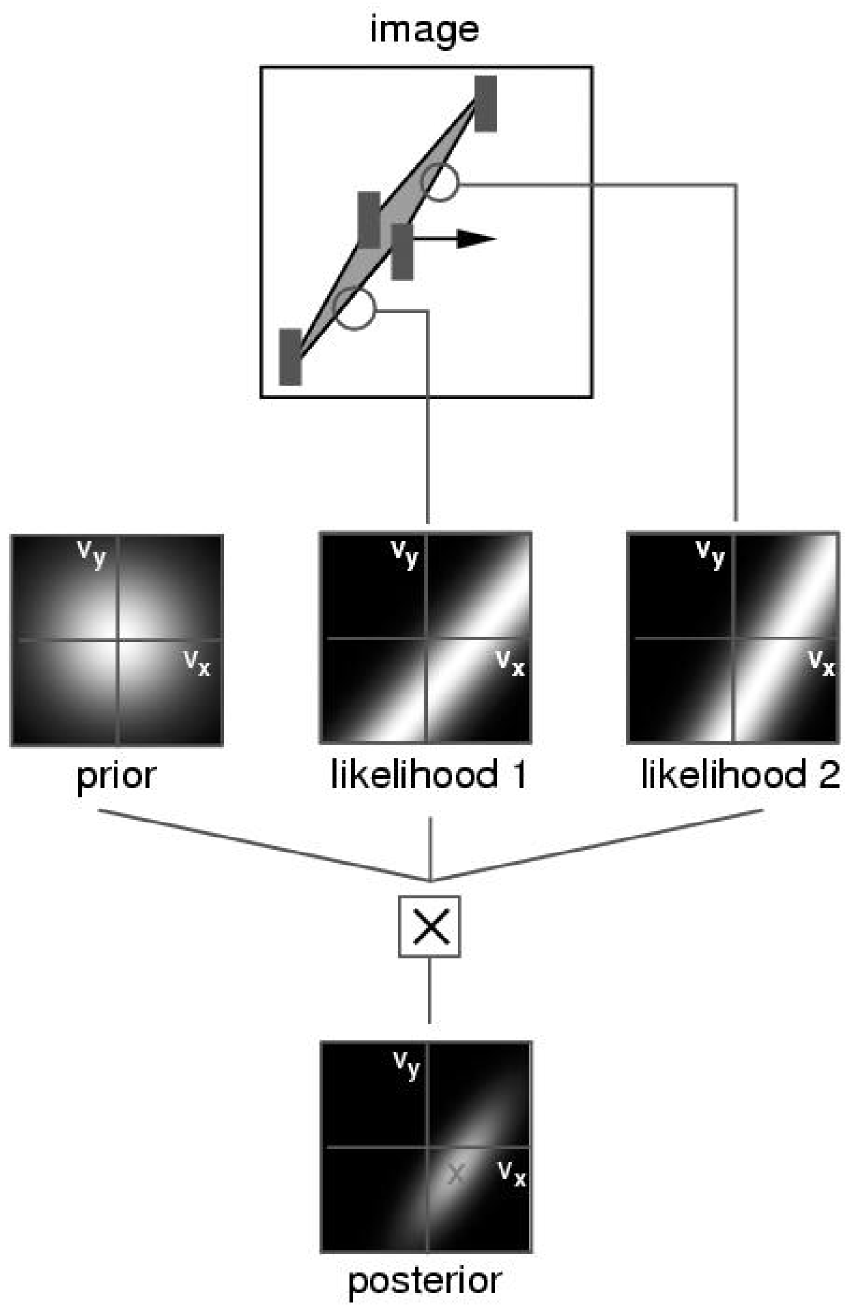
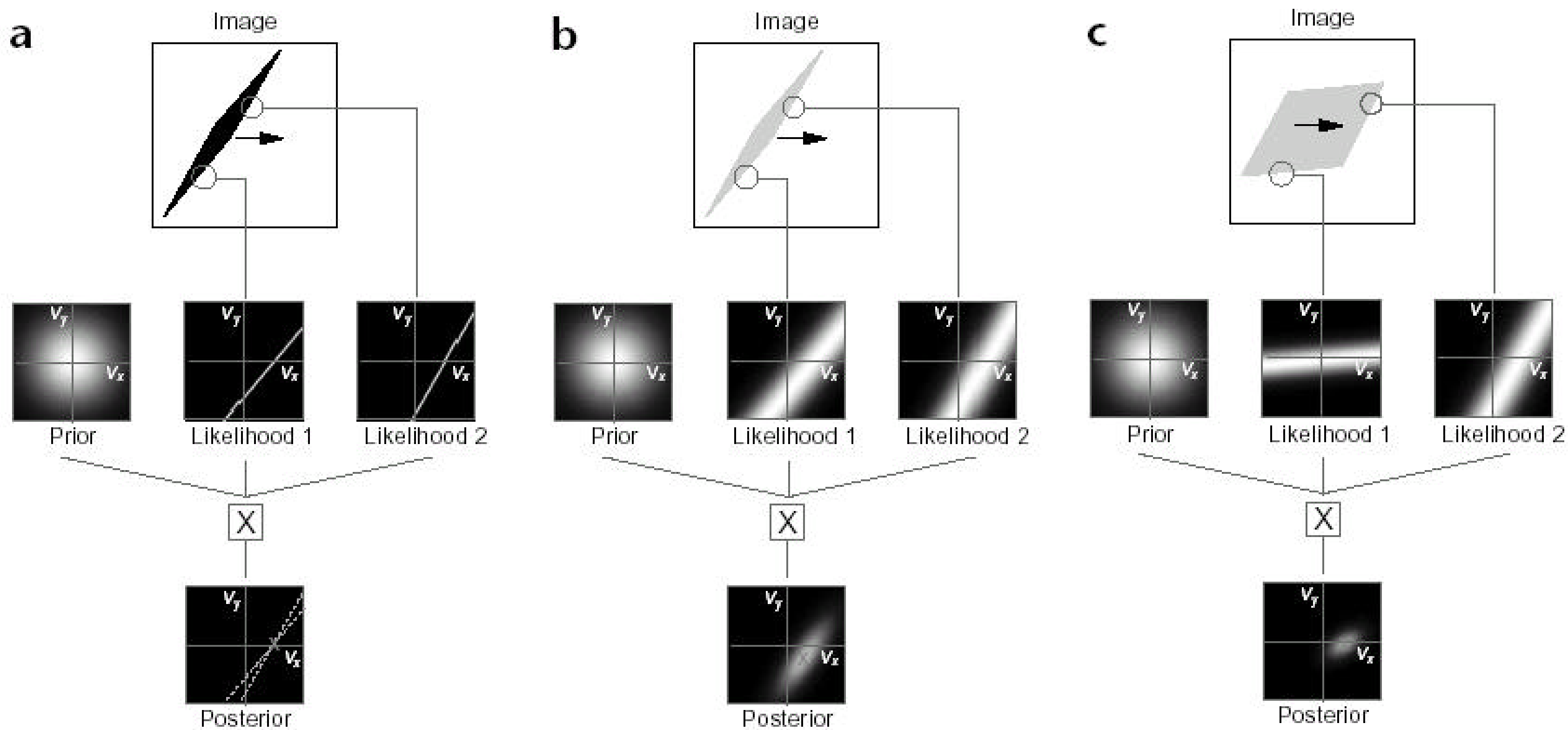


Figure 3: The response of the Bayesian estimator to a narrow rhombus. (replotted from Weiss and Adelson 98)

Effect of contrast



Topics

- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- Direct depth
- SSD tracking
- Robust flow
- Bayesian flow