Visual Computing:

Optical Flow

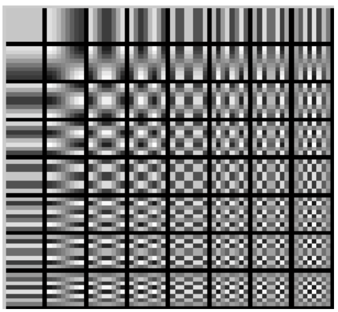
Prof. Marc Pollefeys

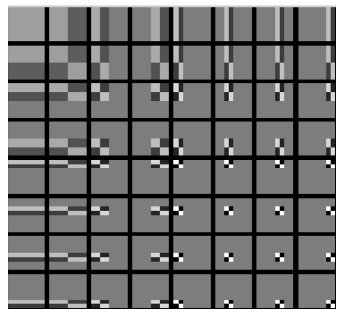
ETH

inf Informatik Computer Science

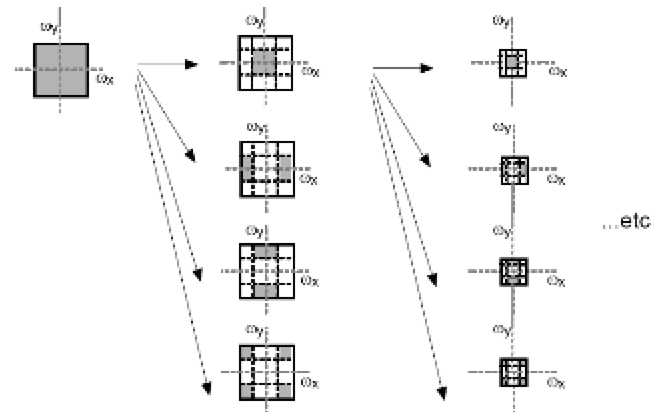
Last lecture

- DCT (JPEG) and DWT (JPEG2000) compression





multi-resolution analysis



- JPEG2000 encoding pipeline

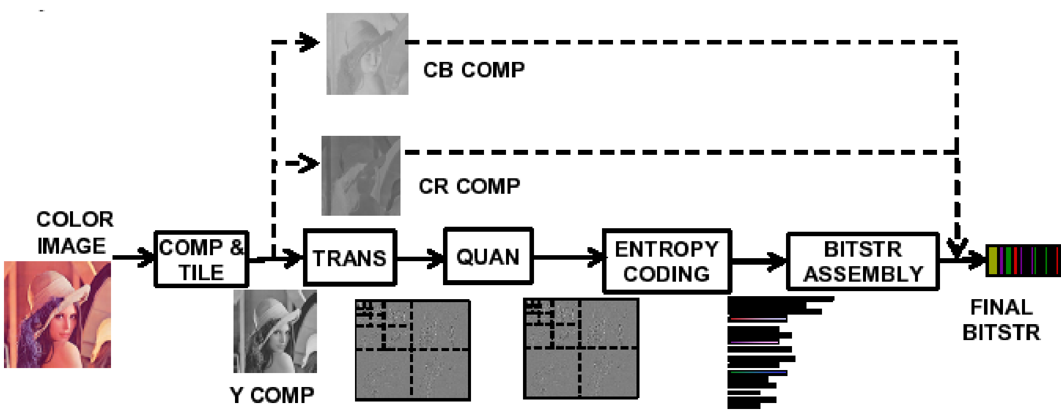
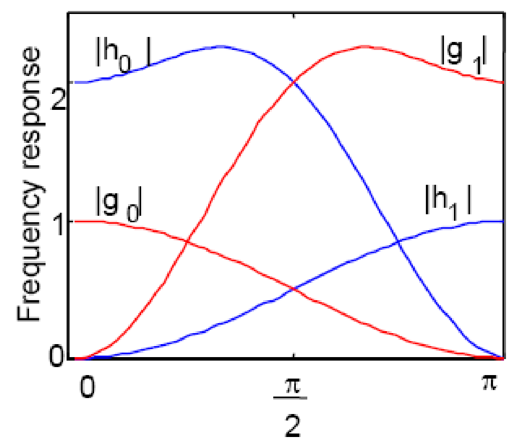


Figure 2 Operation flow of the JPEG 2000 standard.

biortogonal filters

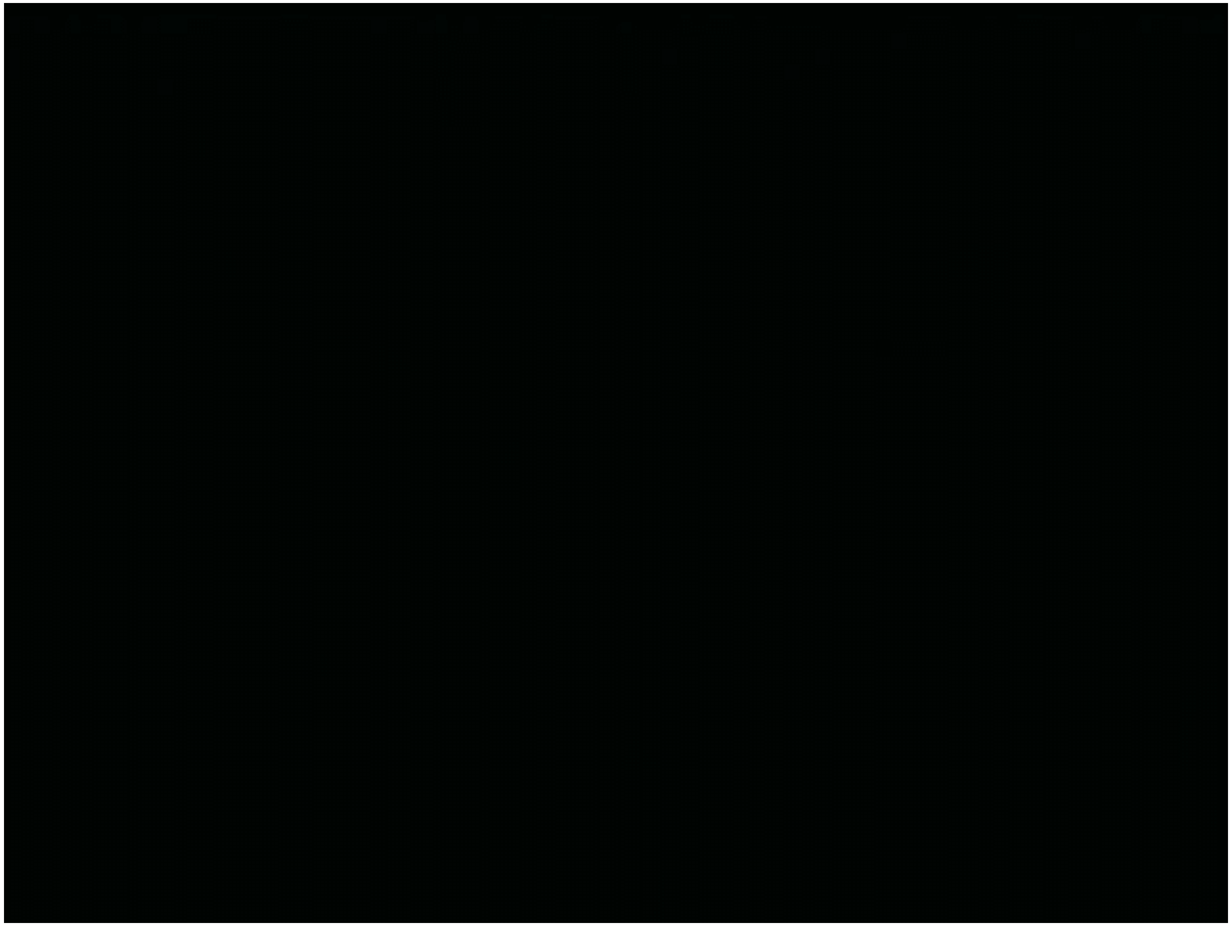


Visual Computing:
Optical Flow

Prof. Marc Pollefeys

ETH

inf | Informatik
Computer Science



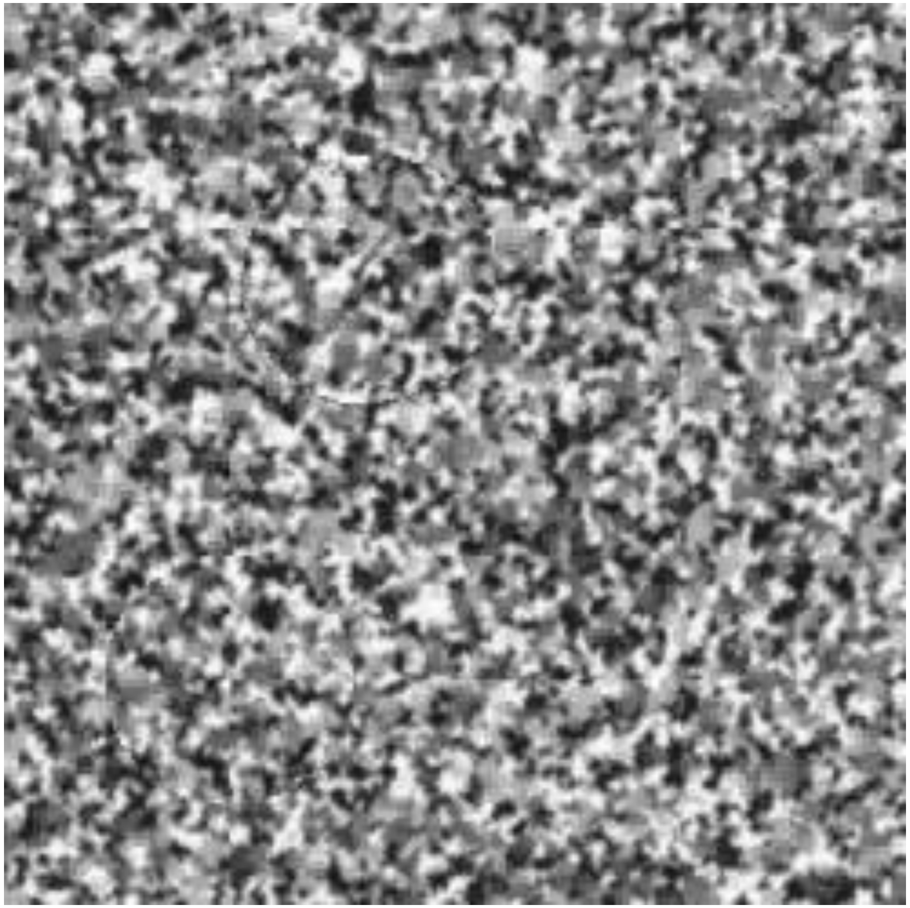
Optical Flow

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

Optical Flow: Where do pixels move to?

Motion is a basic cue

Motion can be the only cue for segmentation



Some slides from Luc Van Gool

Motion is a basic cue

Even impoverished motion data can elicit a strong percept

Applications

- tracking
- structure from motion
- motion segmentation
- stabilization
- compression
- Mosaicing
- …

Optical Flow

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

Definition of Optical Flow

OPTICAL FLOW = apparent motion of
brightness patterns

Ideally, the optical flow is the projection of the three-dimensional velocity vectors on the image

Caution required

Two examples :

1. Uniform, rotating sphere

$\Downarrow$

O.F. = 0

2. No motion, but changing lighting

$\Downarrow$

O.F. $\neq$ 0

Caution required

Mathematical formulation

$I(x,y,t)$ = brightness at (x,y) at time t

Brightness constancy assumption:

$$I\left(x+\frac{dx}{dt}\delta t,\, y+\frac{dy}{dt}\delta t,\, t+\delta t\right)=I(x,y,t)$$

Optical flow constraint equation :

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

Optical Flow

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

The aperture problem

$$u = \frac{dx}{dt}, \qquad v = \frac{dy}{dt}$$

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

$$\boxed{I_x u + I_y v + I_t = 0}$$

1 equation in 2 unknowns

The aperture problem

$$uI_x + vI_y + I_t = 0$$

one equation, two unknowns

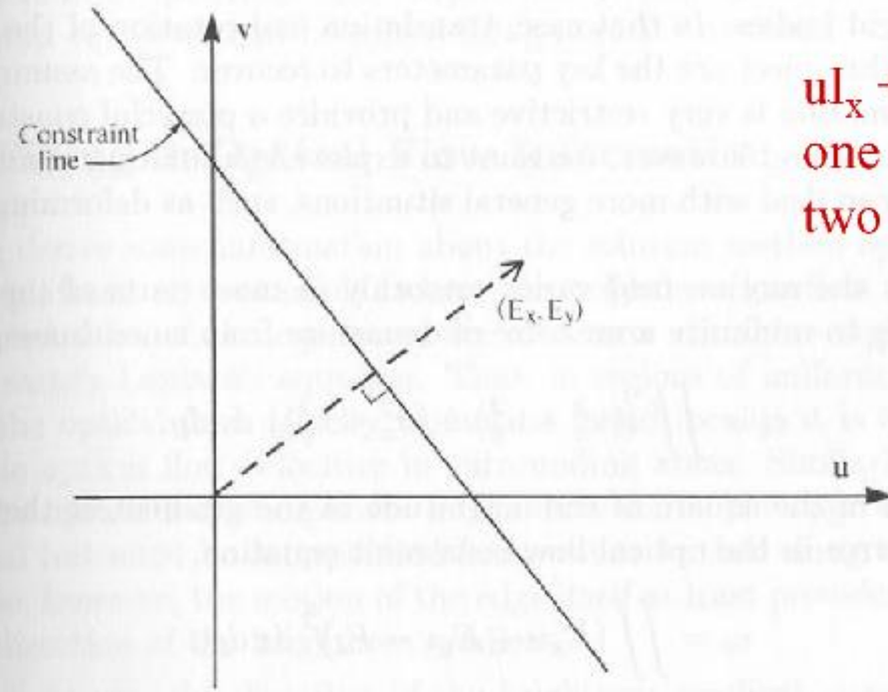
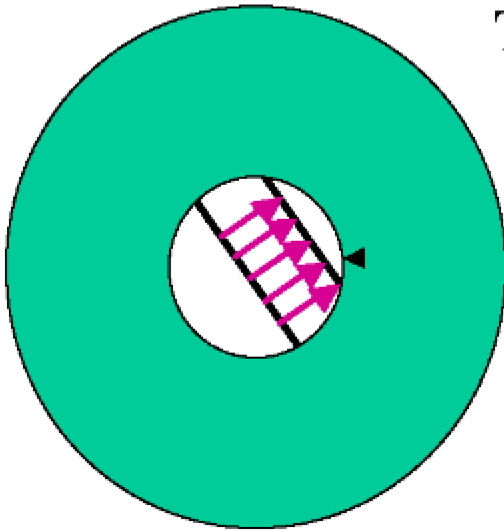


Figure 12-4. Local information on the brightness gradient and the rate of change of brightness with time provides only one constraint on the components of the optical flow vector. The flow velocity has to lie along a straight line perpendicular to the direction of the brightness gradient. We can only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

Aperture problem and Normal Flow

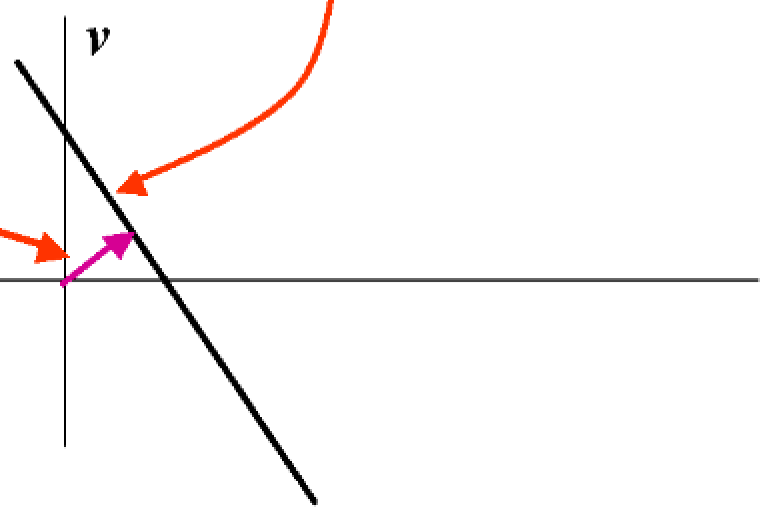


The gradient constraint:

$$I_x u + I_y v + I_t = 0$$

$$\nabla I \bullet \vec{U} = 0$$

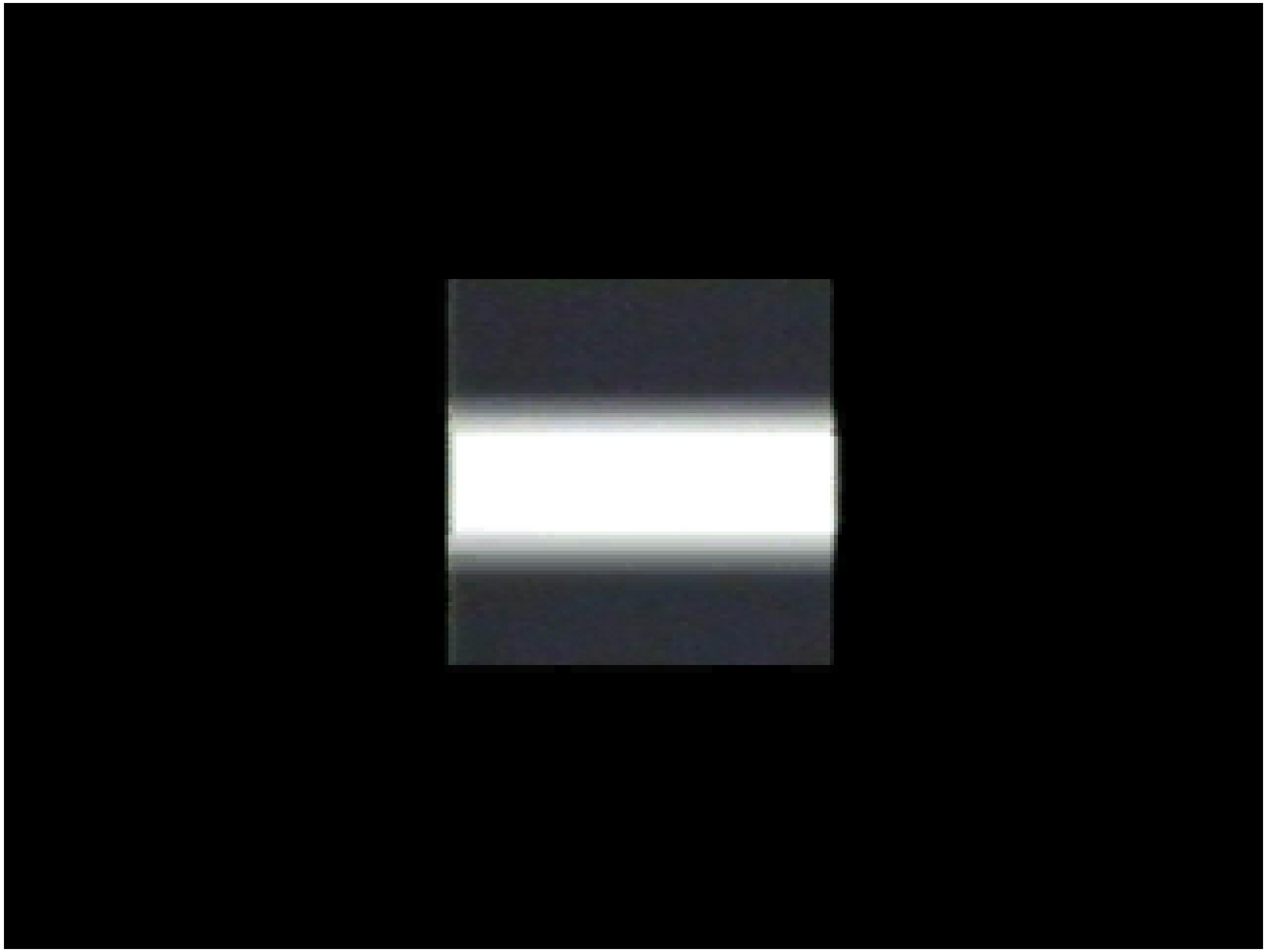
Defines a line in the *(u,v)* space



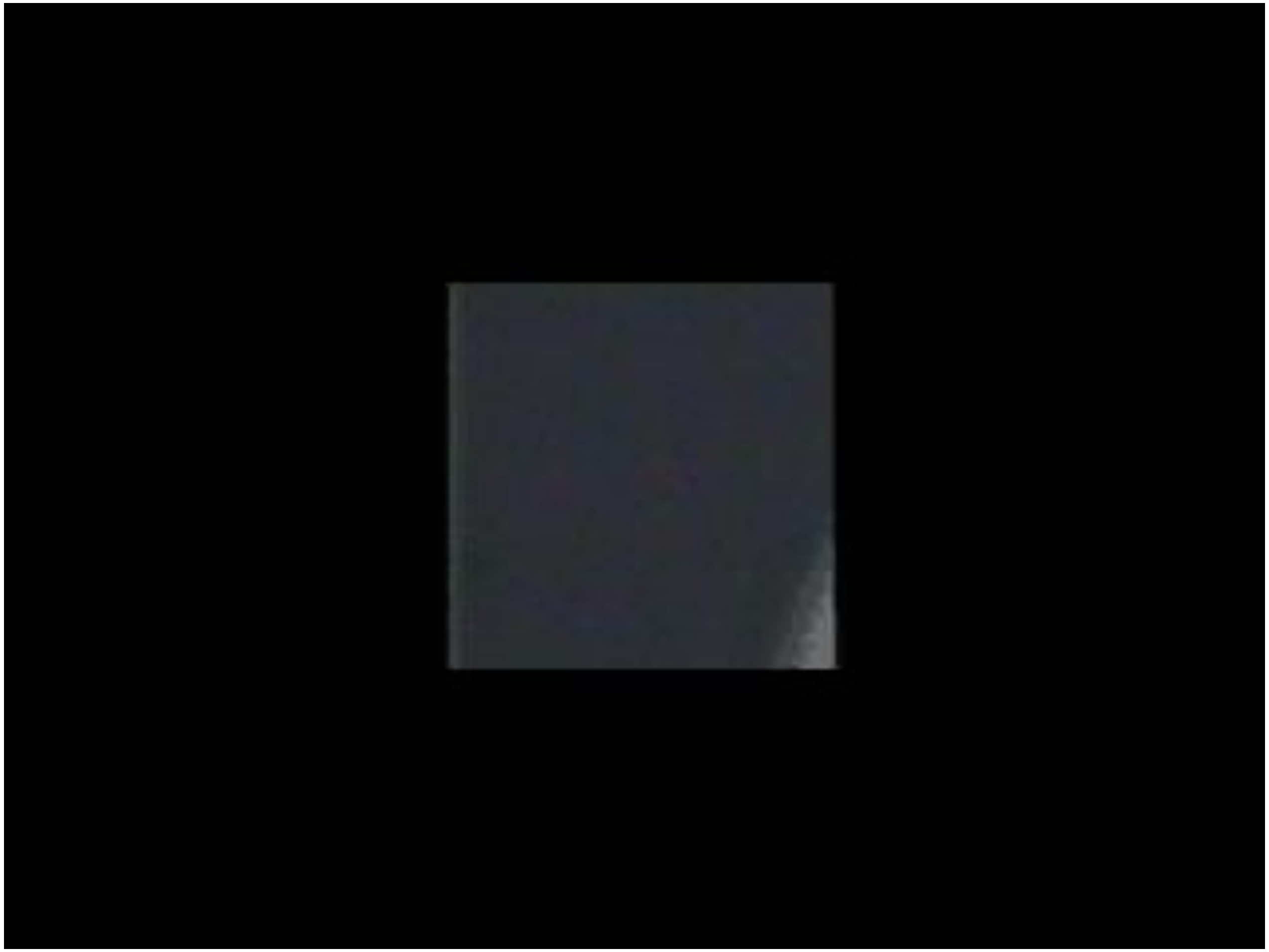
Normal Flow:

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

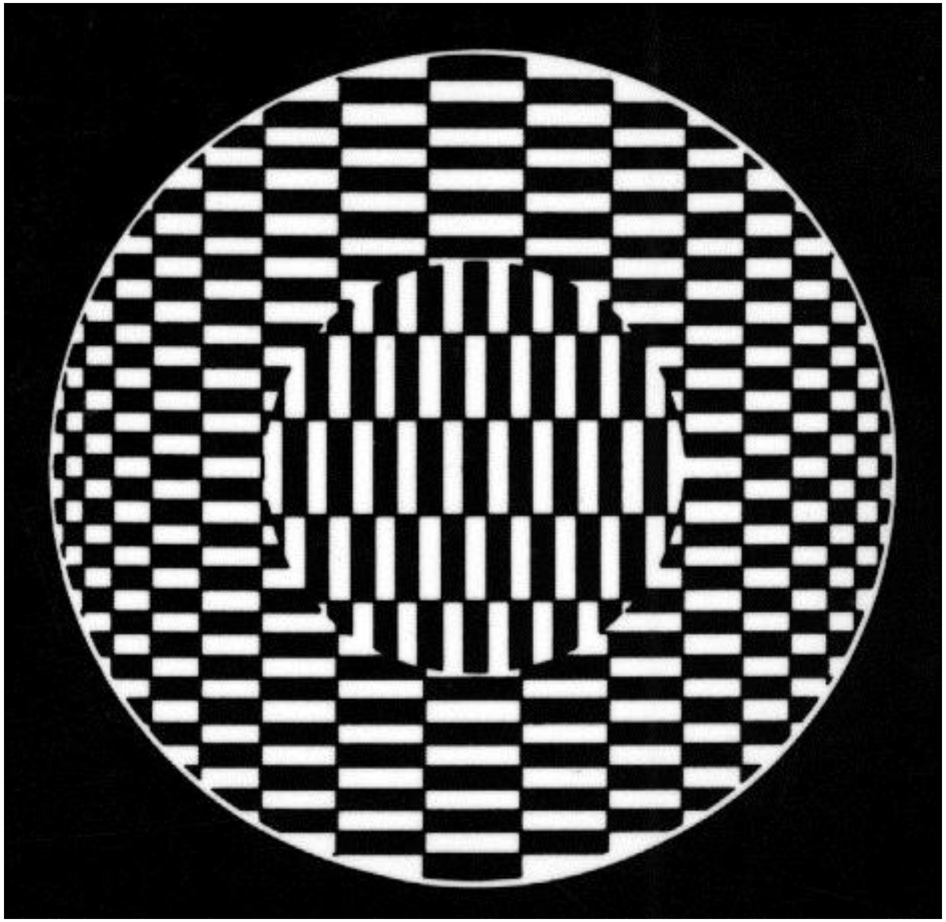
The aperture problem



Remarks



Apparently an aperture problem



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

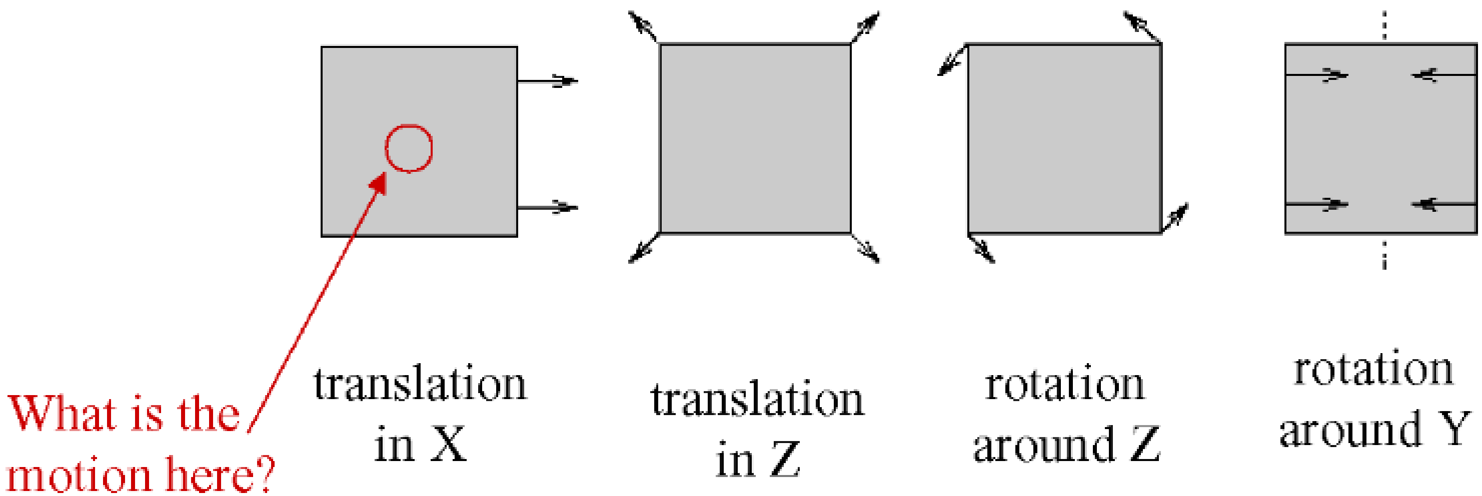
    observed tangent motion
  - Optic Flow

    apparent motion of the brightness pattern

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

Planar motion examples

- Ideal motion of a plane



Scene Flow: →

Normal Flow: undef

Optic Flow: ?, probably 0

Optical Flow

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

Horn & Schunck algorithm

Additional smoothness constraint :

$$e_s = \iint ((u_x^2 + u_y^2) + (v_x^2 + v_y^2))dxdy,$$

besides OF constraint equation term

$$e_c = \iint (I_x u + I_y v + I_t)^2 dxdy,$$

minimize $e_s + \lambda e_c$

ETH

Horn & Schunck

The Euler-Lagrange equations :

$$F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} = 0$$

$$F_v - \frac{\partial}{\partial x} F_{v_x} - \frac{\partial}{\partial y} F_{v_y} = 0$$

In our case ,

$$F = (u_x^2 + u_y^2) + (v_x^2 + v_y^2) + \lambda (I_x u + I_y v + I_t)^2,$$

so the Euler-Lagrange equations are

$$\Delta u = \lambda (I_x u + I_y v + I_t) I_x,$$

$$\Delta v = \lambda (I_x u + I_y v + I_t) I_y,$$

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$ is the Laplacian operator

Horn & Schunk

Remarks :

1. Coupled PDEs solved using iterative methods and finite differences

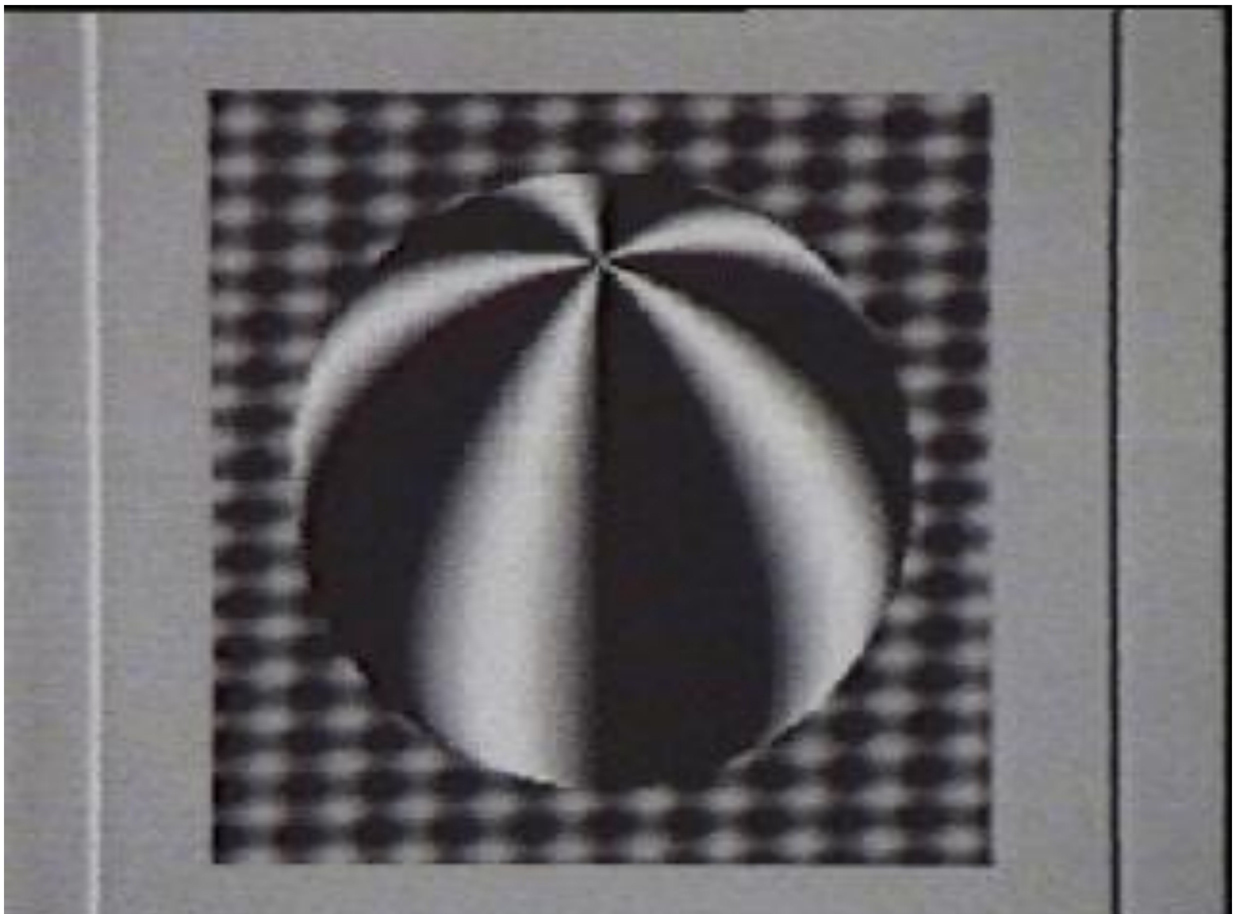
$$\frac{\partial u}{\partial t} = \Delta u - \lambda(I_x u + I_y v + I_t) I_x ,$$

$$\frac{\partial v}{\partial t} = \Delta v - \lambda(I_x u + I_y v + I_t) I_y ,$$

2. More than two frames allow a better estimation of I_t

3. Information spreads from corner-type patterns

ETH

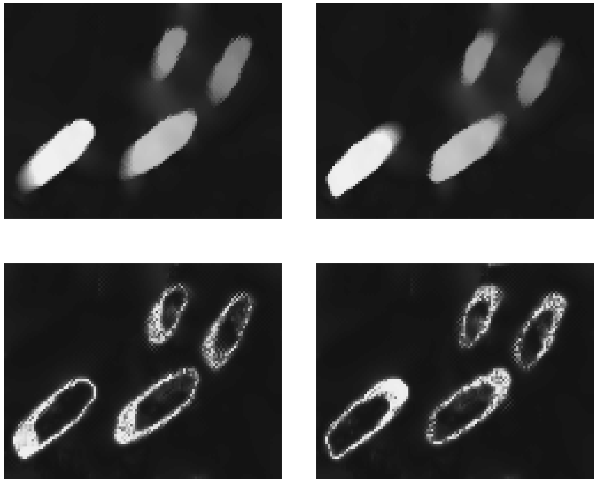


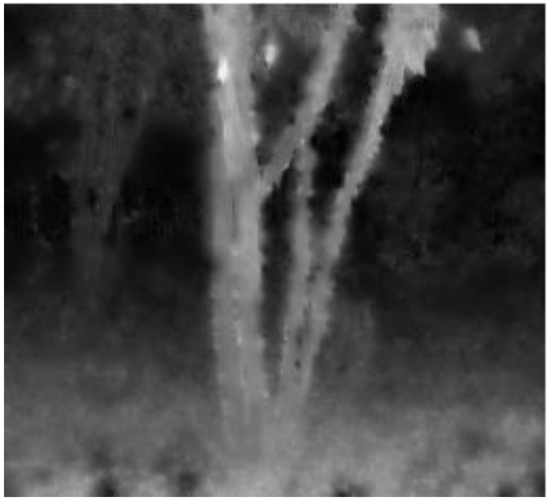
Horn & Schunk, remarks

1. Errors at boundaries

2. Example of *regularisation*
   (selection principle for the solution of illposed problems)

Results of an enhanced system





Structure from motion with OF

Optical Flow

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

Lucas-Kanade: Integrate over a Patch

Assume a single velocity for all pixels within an image patch

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

$$\frac{dE(u,v)}{du} = \sum 2I_x\left(I_x u + I_y v + I_t\right) = 0$$

$$\frac{dE(u,v)}{dv} = \sum 2I_y\left(I_x u + I_y v + I_t\right) = 0$$

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

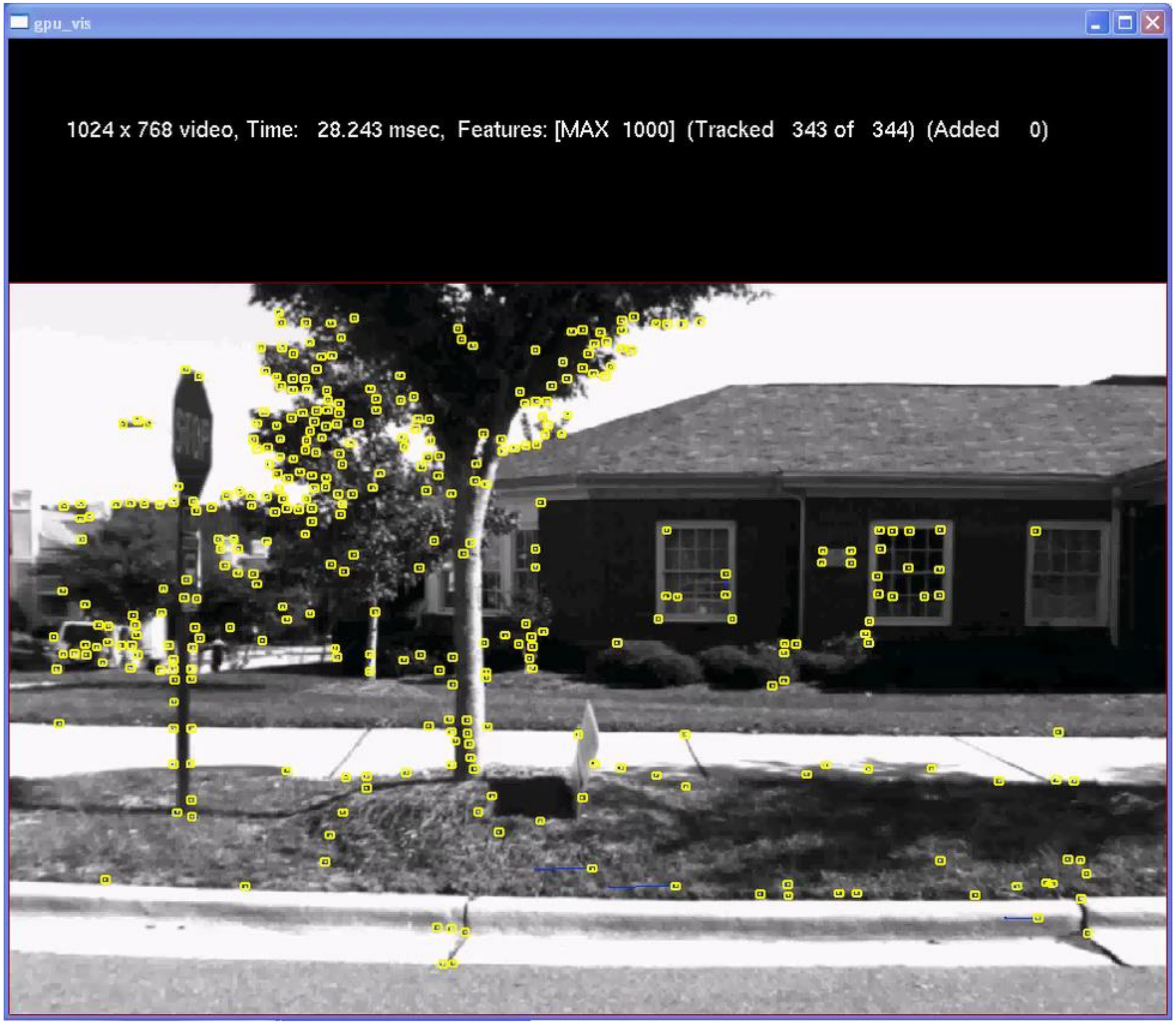
Let $M = \sum (\nabla I)(\nabla I)^T$ and $b = \begin{bmatrix} -\sum I_x I_t \\ -\sum I_y I_t \end{bmatrix}$

- 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

KLT feature tracker:
see "Good Features to Track", *Shi and Tomasi, CVPR'94*, 1994, pp. 593 - 600.


gpu_vis
1024 x 768 video, Time: 28.243 msec, Features: [MAX 1000] (Tracked 343 of 344) (Added 0)
STOP

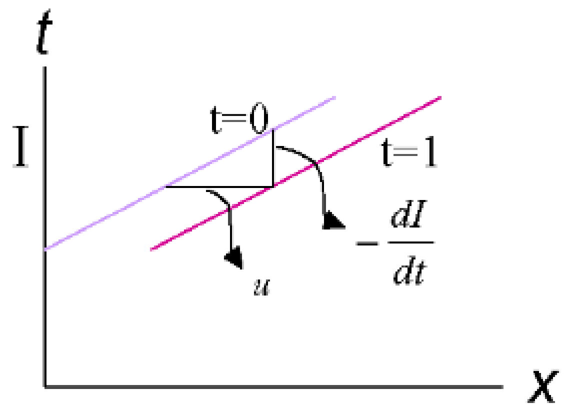
Iterative Refinement

- 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



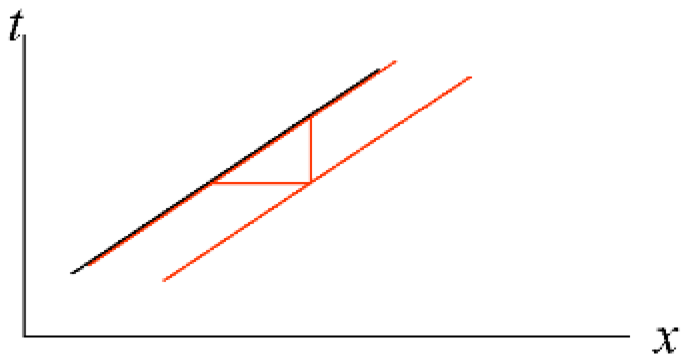
$$\frac{dI}{dx} = -\frac{dI/dt}{u}$$

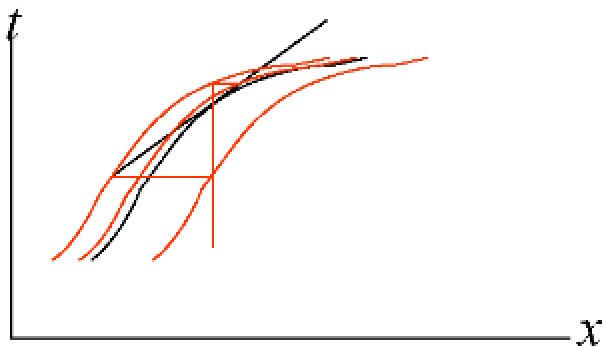
$$0 = I_x u + I_t$$

"shift by u to account for I_x with I_t"

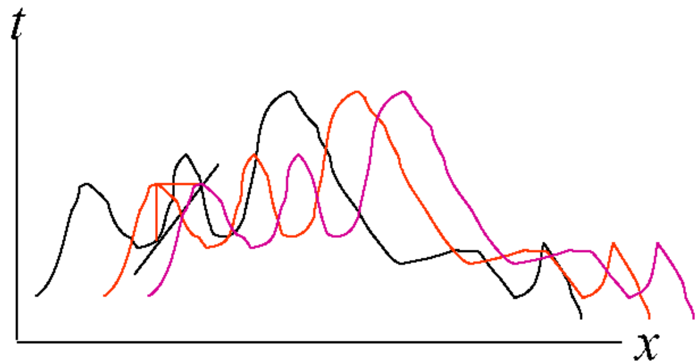
$$u = -\frac{I_t}{I_x}$$

Iterative refinement





BUT!!



Optical Flow

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

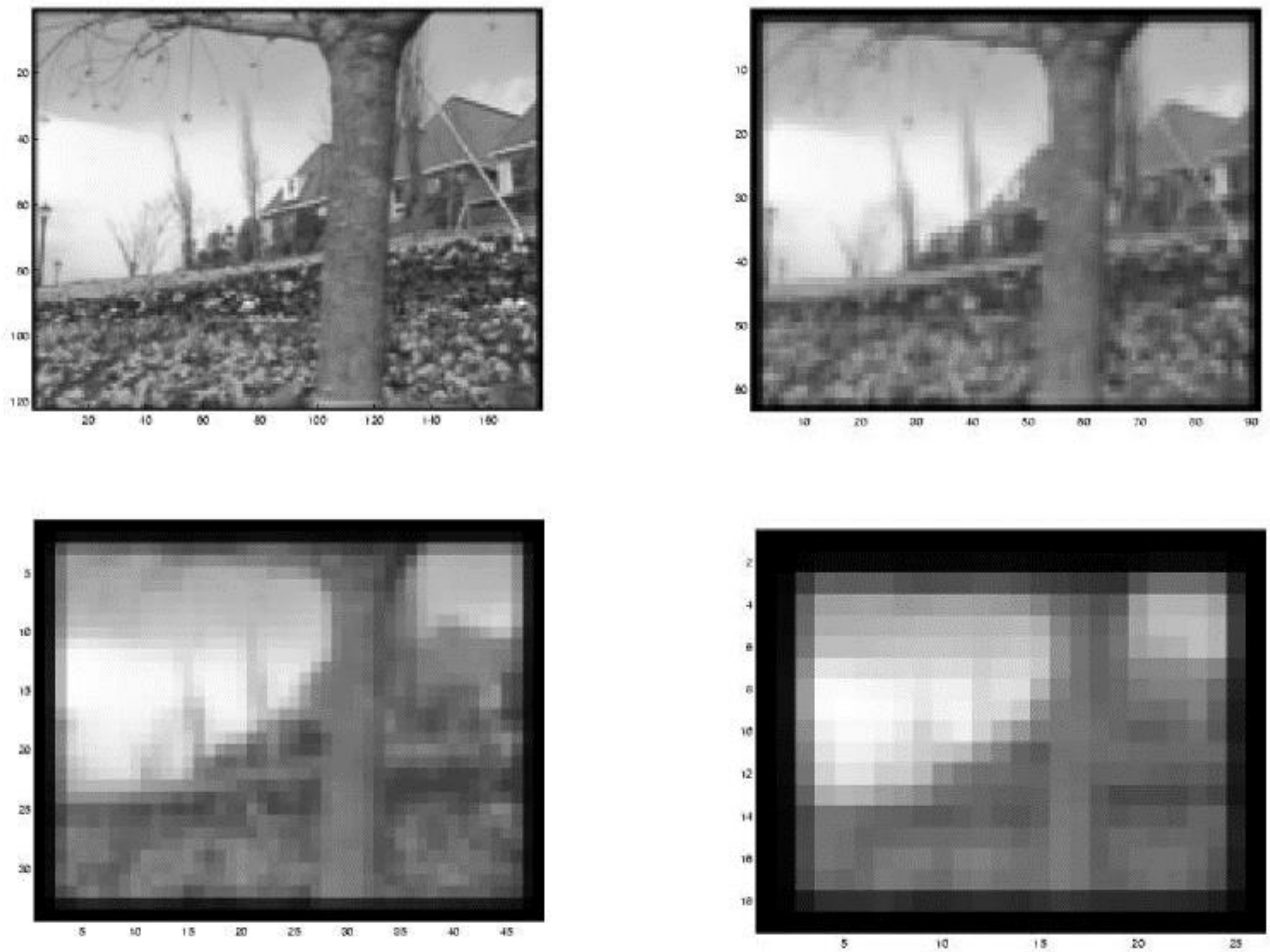
Limits of the (local) gradient method

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*

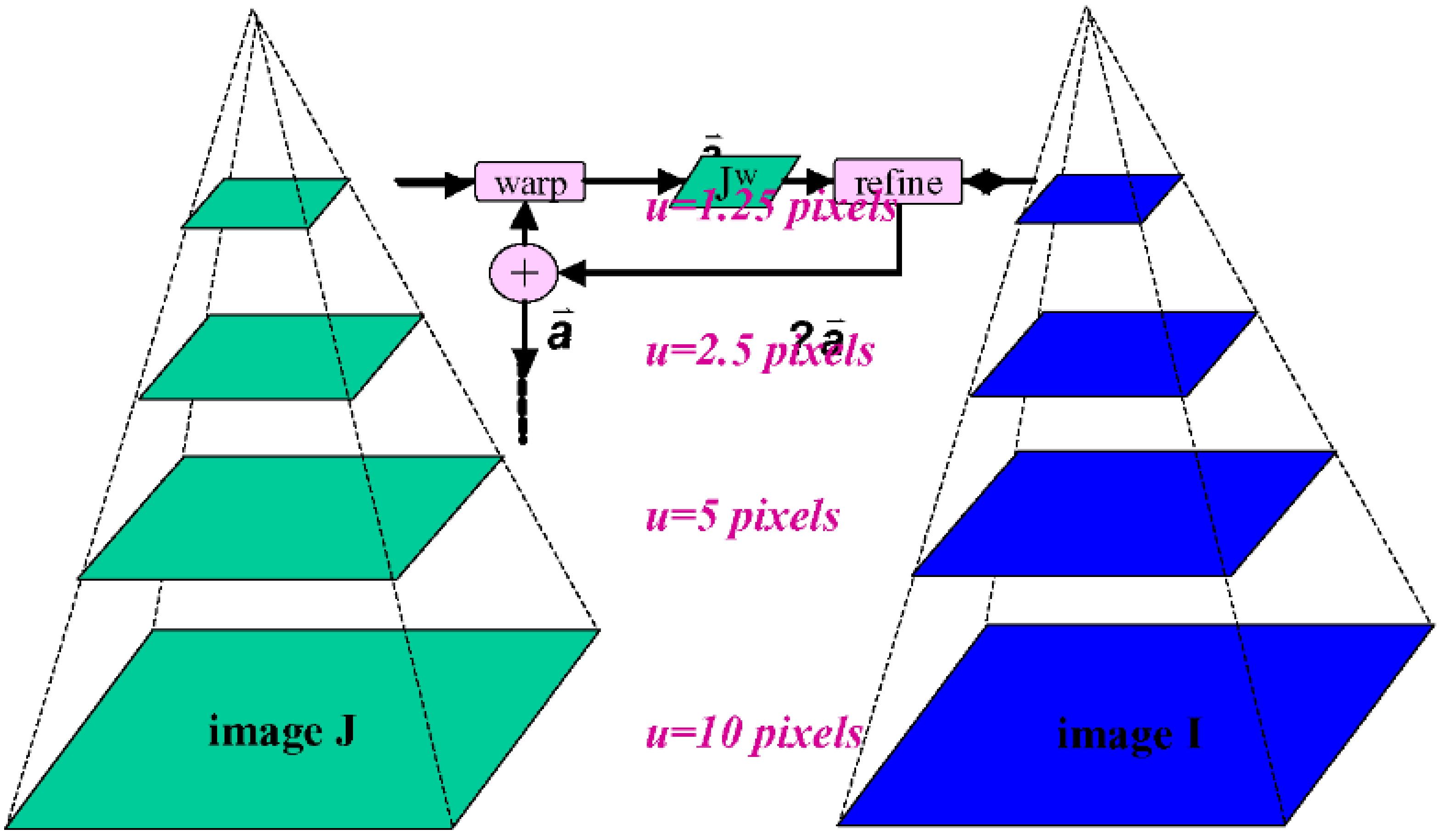
Pyramid / "Coarse-to-fine"



Coarse-to-Fine Estimation

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

==> small *u* and *v* ...



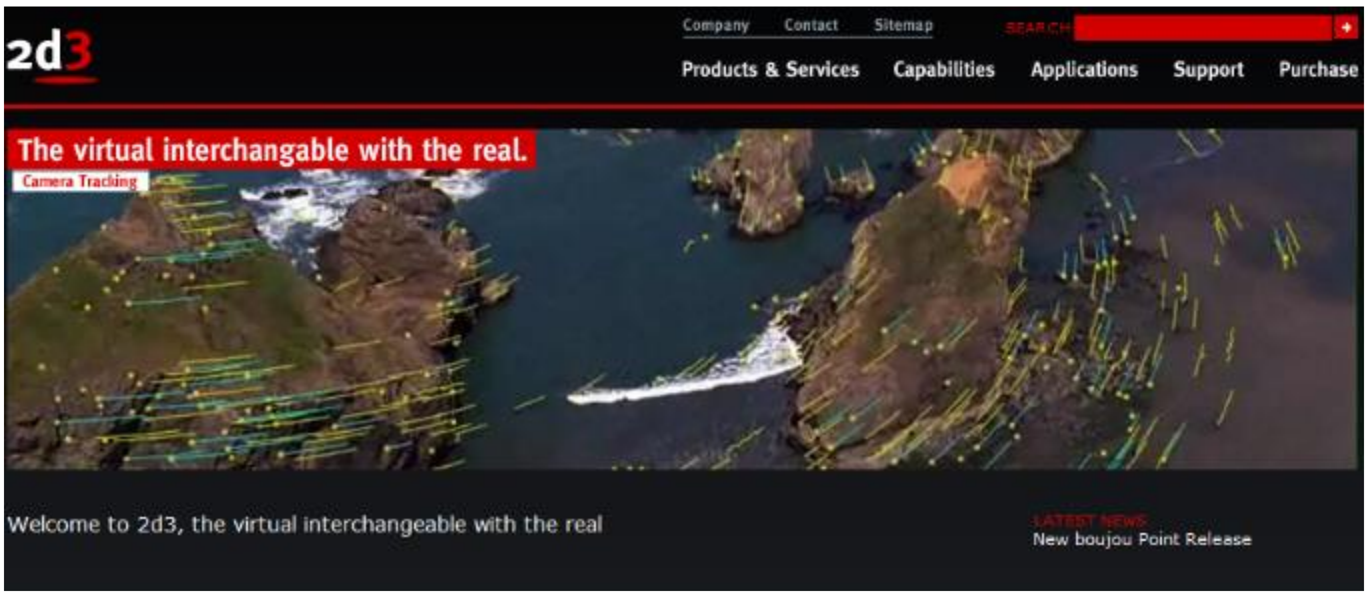
Pyramid of image J

Pyramid of image I

OF application: Image stabilization

DeShaker

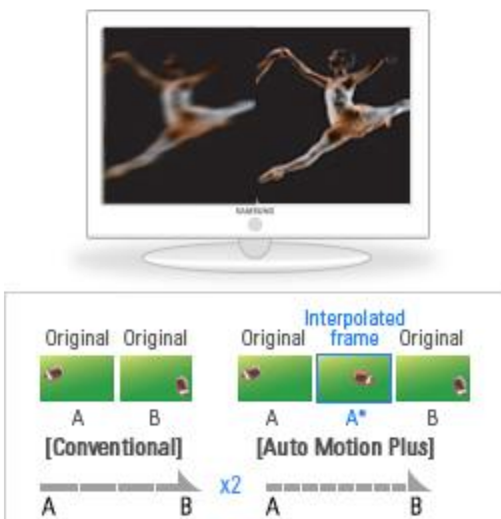
OF application: MatchMoving

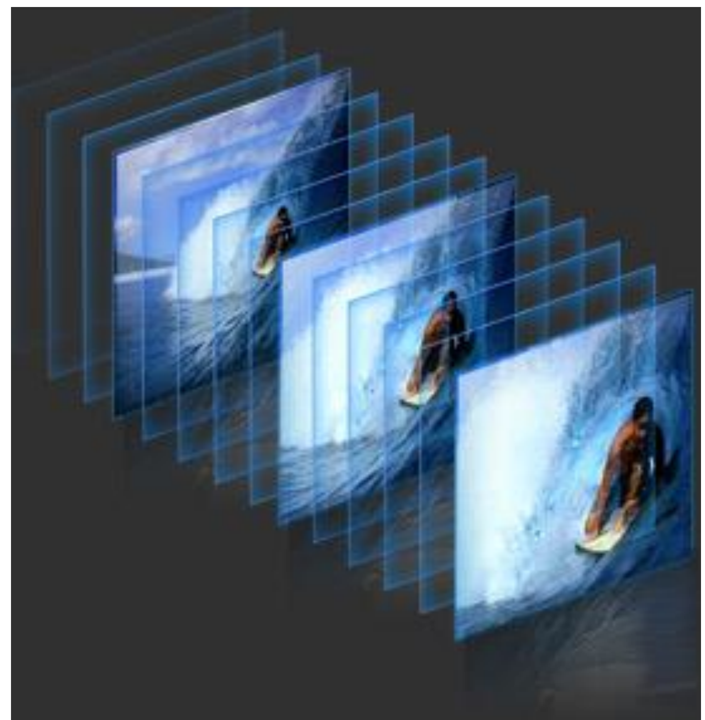



boujou2

OF application: Slow motion

- Slow motion (generate intermediate frames)
- Technology is also key to 100Hz television

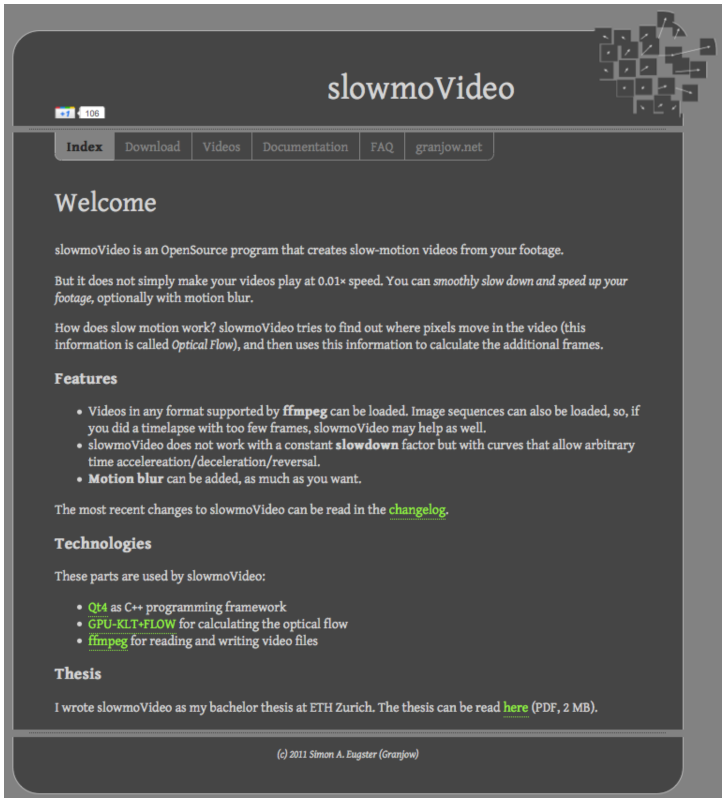




ETH

SlowMoVideo

Bachelor thesis Simon Eugster



http://slowmovideo.granjow.net/

ETH

Optical Flow

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

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: $\mathbf{h} = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$

What about other types of motion?

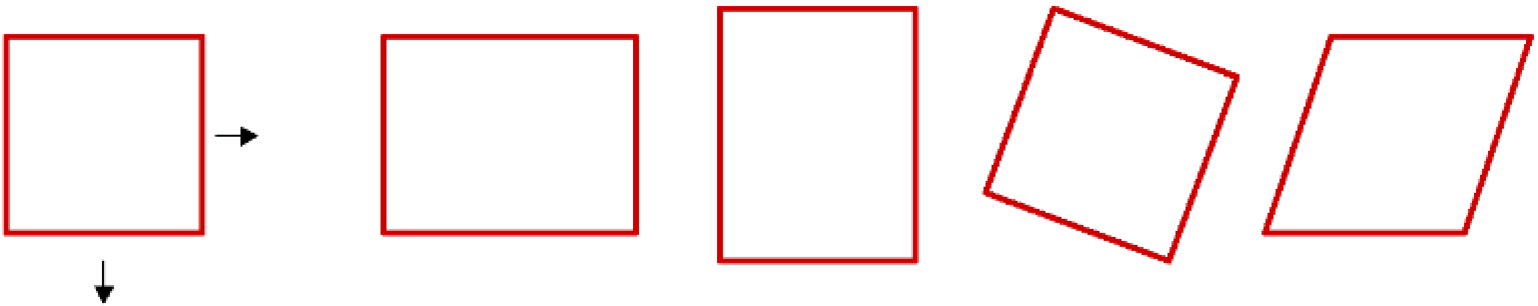
Generalization

- 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: $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \mathbf{h} = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$

Generalization



Affine: $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\mathbf{h} = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$

Example: Affine Motion

$$u(x,y) = a_1 + a_2 x + a_3 y$$
$$v(x,y) = a_4 + a_5 x + a_6 y$$

Substituting into the B.C. Equation:

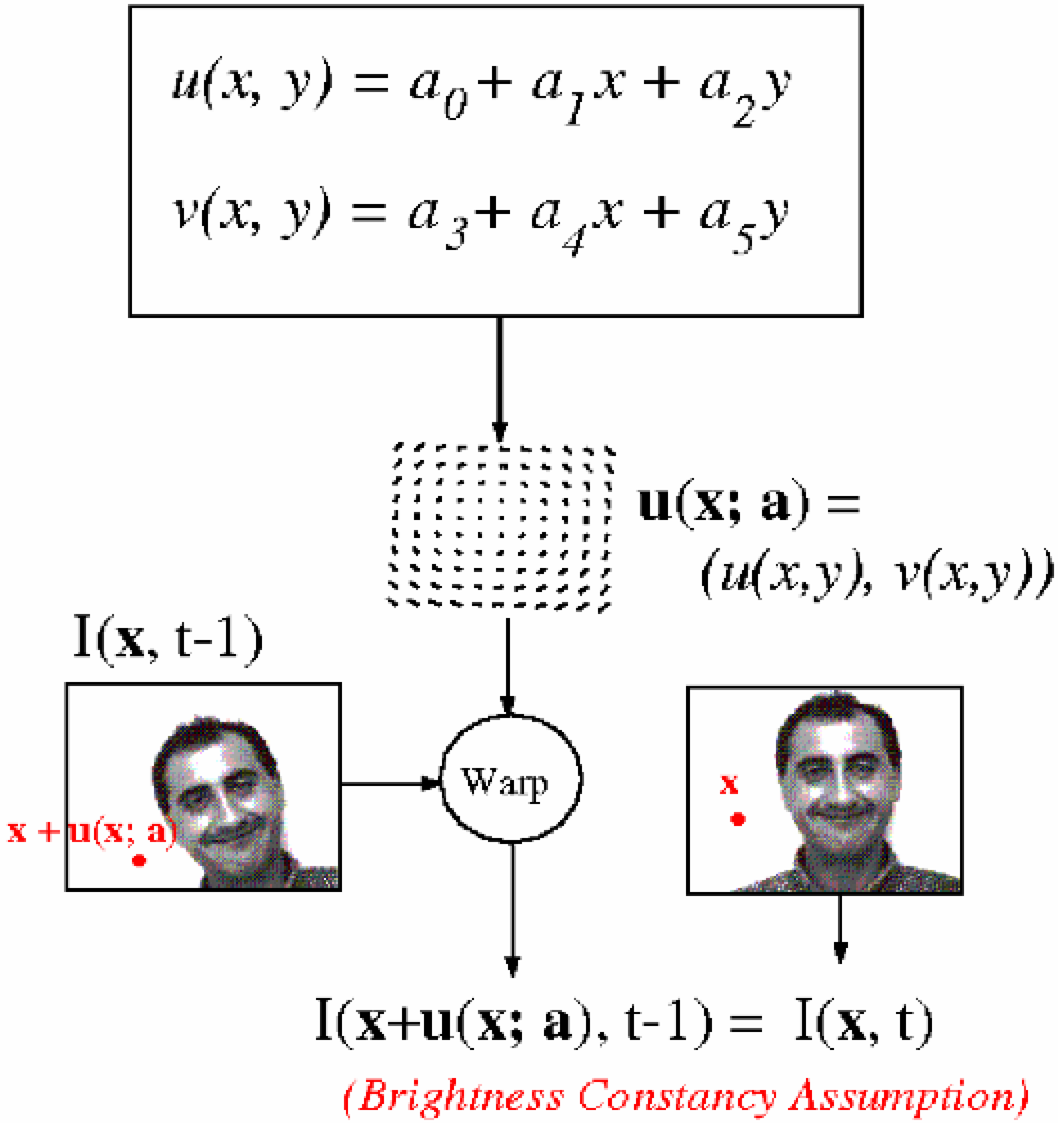
$$I_x \cdot u + I_y \cdot v + I_t \approx 0$$

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

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

Least Square Minimization (over all pixels):

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



$u(x, y) = a_0 + a_1 x + a_2 y$
$v(x, y) = a_3 + a_4 x + a_5 y$
$\mathbf{u}(\mathbf{x}; \mathbf{a}) =$
$(u(x,y), v(x,y))$
I(**x**, t-1)
x + u(x; a)
Warp
x
I(**x**+**u**(**x**; **a**), t-1) = I(**x**, t)
(Brightness Constancy Assumption)

KLT: Good features to keep tracking

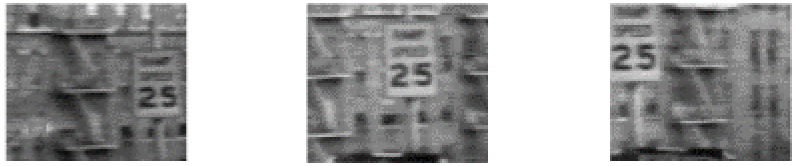
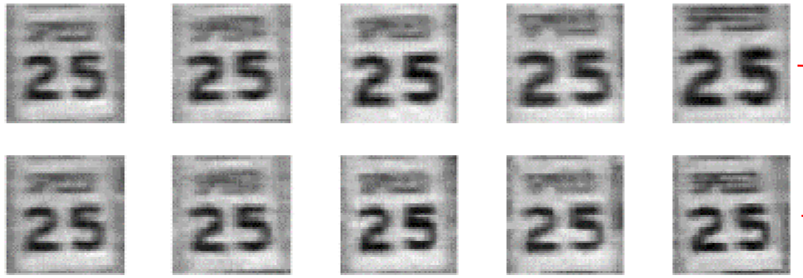
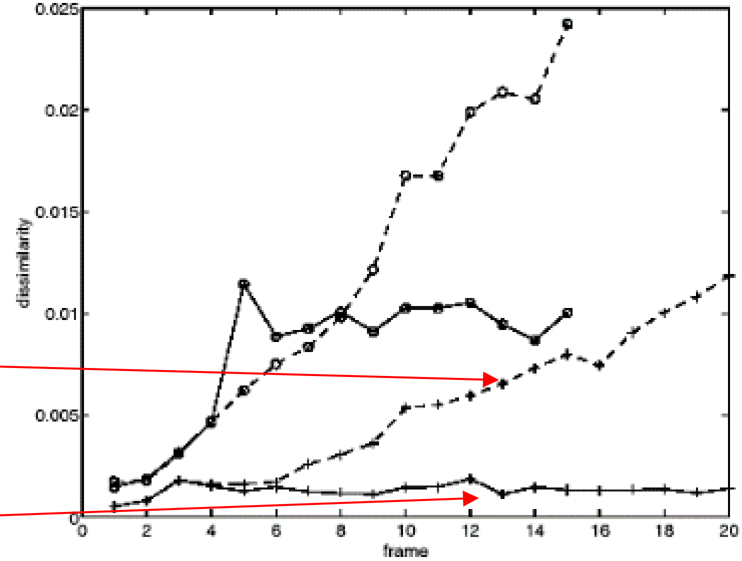


Figure 1: Three frame details from Woody Allen's *Manhattan*. The details are from the 1st, 11th, and 21st frames of a subsequence from the movie.





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

Generalization

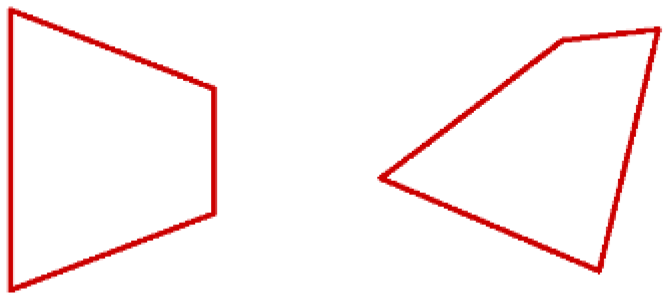
- Transformations/warping of image

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

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 +



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

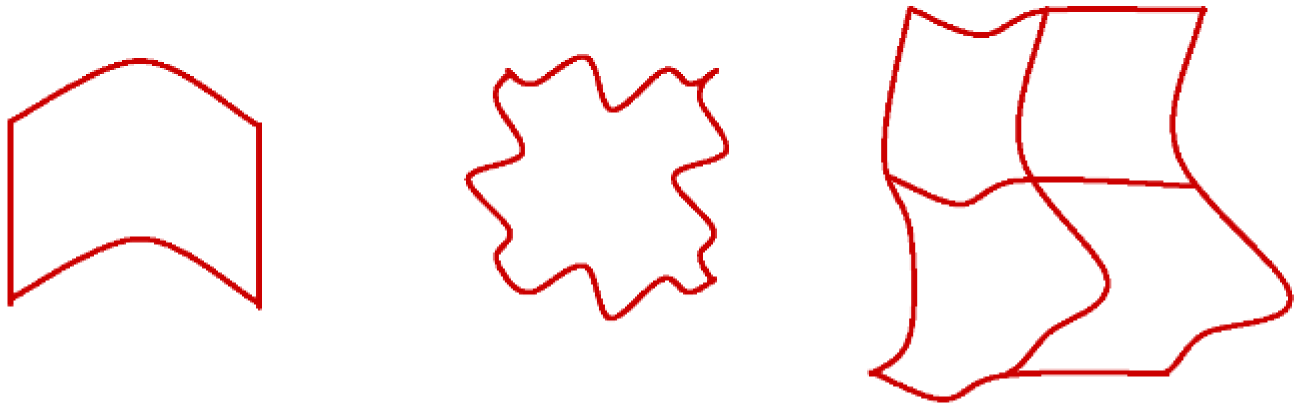
Generalization

- 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

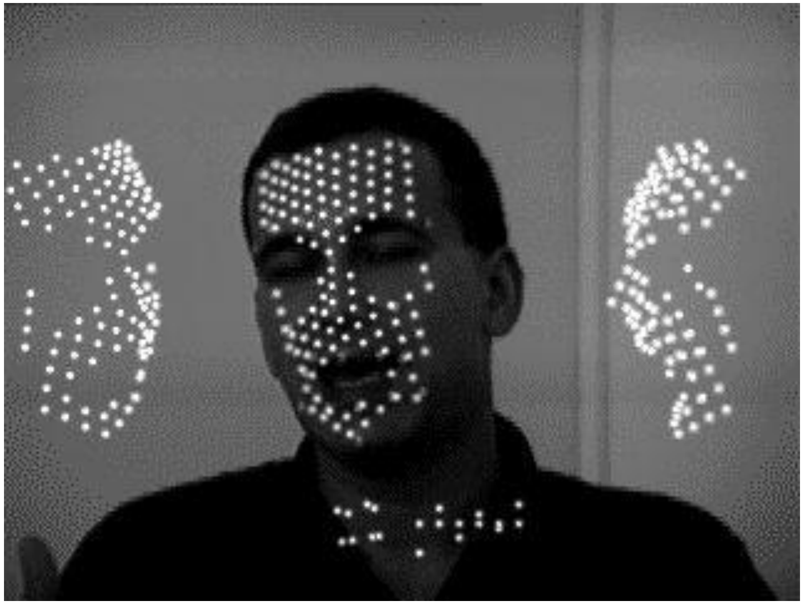
$$x' = \frac{h_1 + h_2 x + h_3 y}{h_7 + h_8 x + h_9 y}$$
$$y' = \frac{h_4 + h_5 x + h_6 y}{h_7 + h_8 x + h_9 y}$$

and

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

Advanced parametric model

- Optical flow constrained by non-rigid face model



Flexible flow for 3D nonrigid tracking and shape recovery,
Brand and Bhotika, CVPR2001.

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: $\gamma(x, y)$

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

Residual Planar Parallax Motion

Global parameters: t_1, t_2, t_3

Local Parameter: $\gamma(x, y)$

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

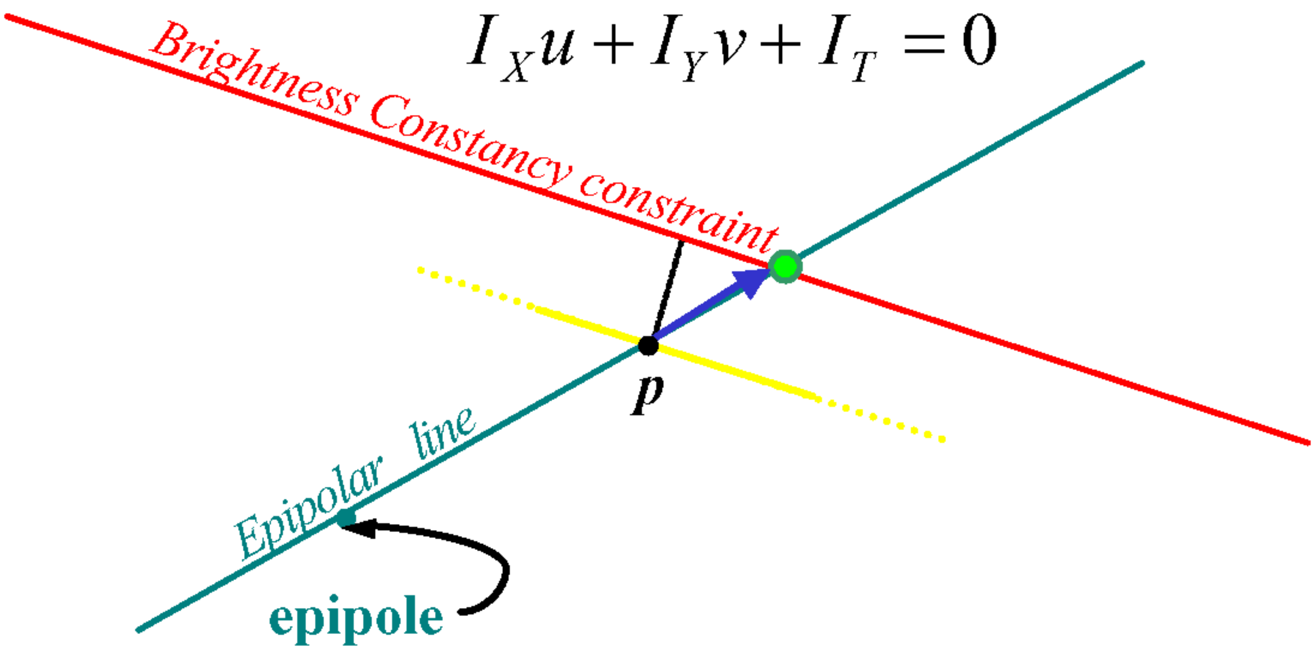
Residual Planar Parallax Motion
(Plane+Parallax)

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.

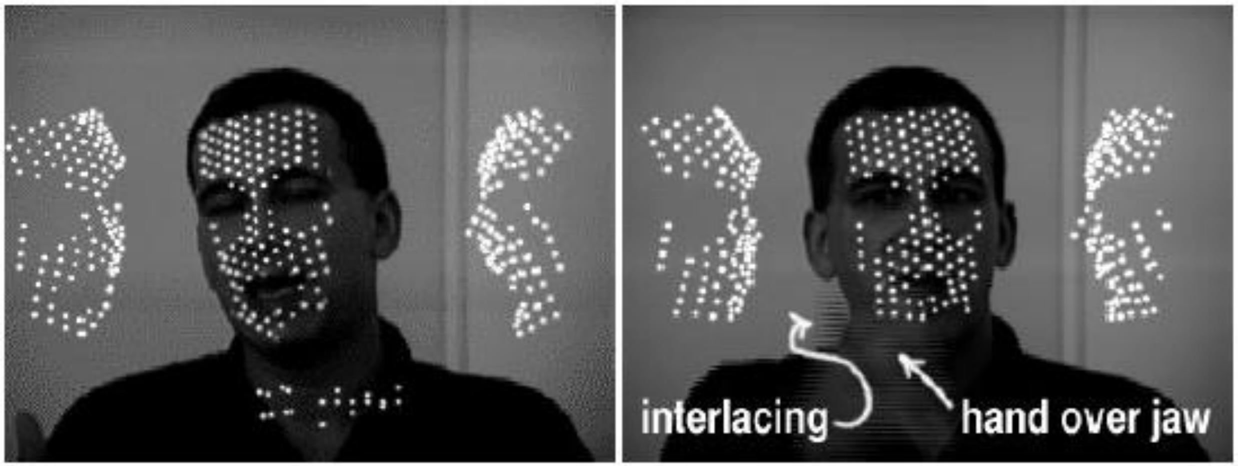


Figure 1: Model-based tracking is robust to degraded images and transient occlusions. Dots show flexed model in 3/4, frontal, and profile view. Dots on face show where the image is sampled. Dots on neck encode 3D motion parameters.

ETH

Brand, M.E.; Bhotika, R., "Flexible Flow for 3D Nonrigid Tracking and Shape Recovery", *IEEE Computer Society Conference on Computer vision and Pattern Recognition (CVPR)*, ISSN: 1063-6919, Vol. 1, pp. 315-322, December 2001

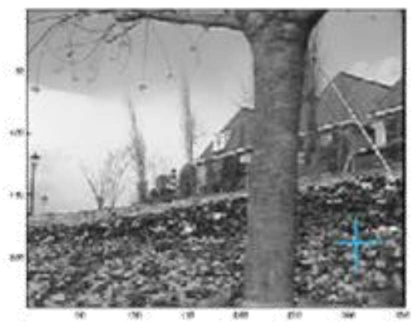
Optical Flow

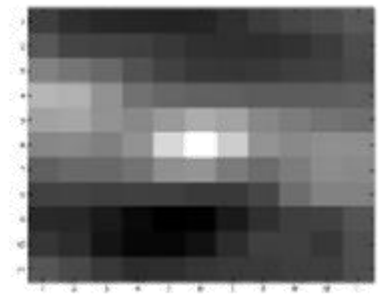
- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- SSD tracking
- Bayesian flow

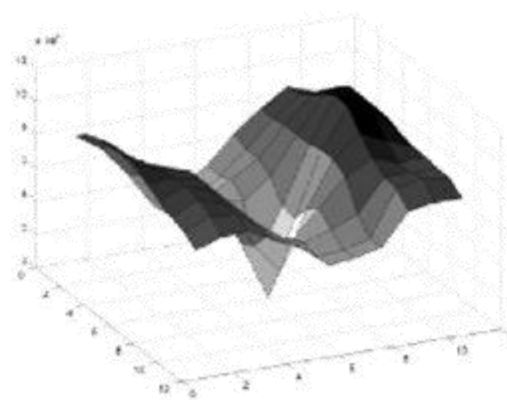
Correlation and SSD

- 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

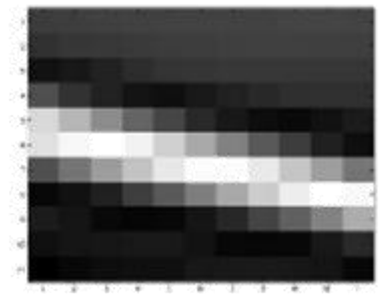
SSD Surface – Textured area

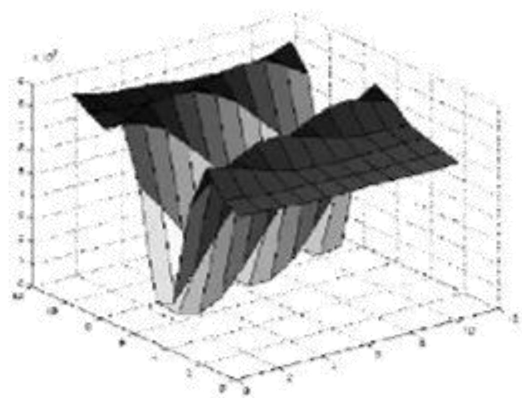




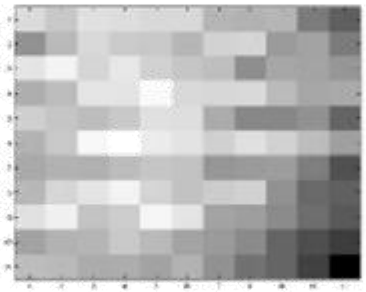


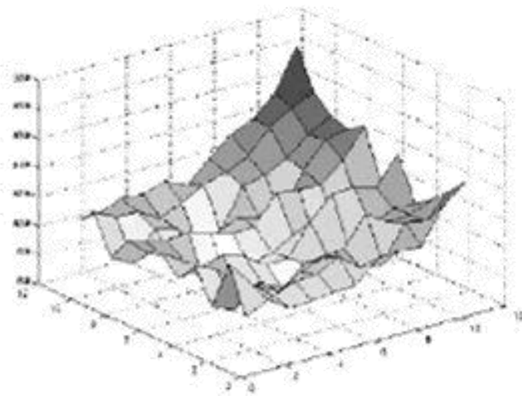
SSD Surface -- Edge





SSD Surface – homogeneous area





Discrete Search vs. Gradient Based Estimation

Consider image I translated by u_0, v_0

$$I_0(x, y) = I(x, y)$$
$$I_1(x + u_0, y + v_0) = I(x, y) + \eta_1(x, y)$$

$$E(u, v) = \sum_{x,y} (I(x, y) - I_1(x + u, y + v))^2$$
$$= \sum_{x,y} (I(x, y) - I(x - u_0 + u, y - v_0 + v) - \eta_1(x, y))^2$$

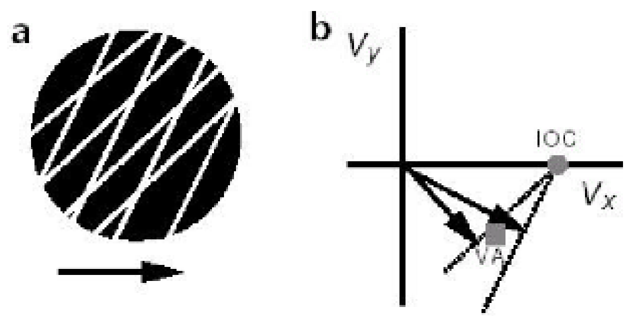
Discrete search simply searches for the best estimate.
Gradient method linearizes the intensity function and solves for the estimate

Optical Flow

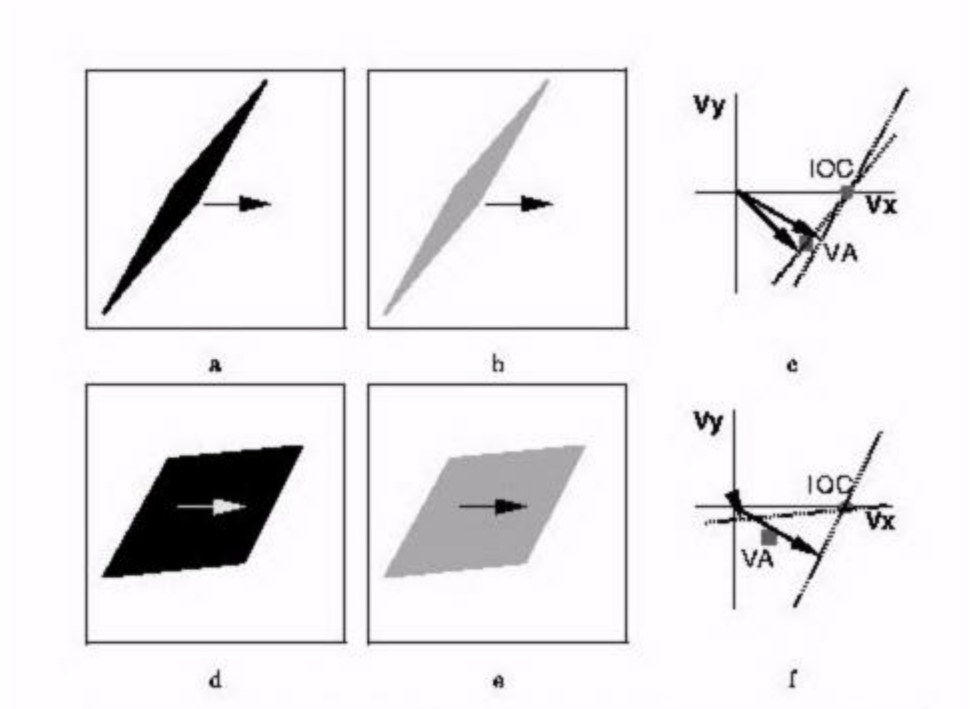
- Brightness Constancy
- The Aperture problem
- Regularization
- Lucas-Kanade
- Coarse-to-fine
- Parametric motion models
- SSD tracking
- Bayesian flow

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):



Rhombus Displays



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

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….

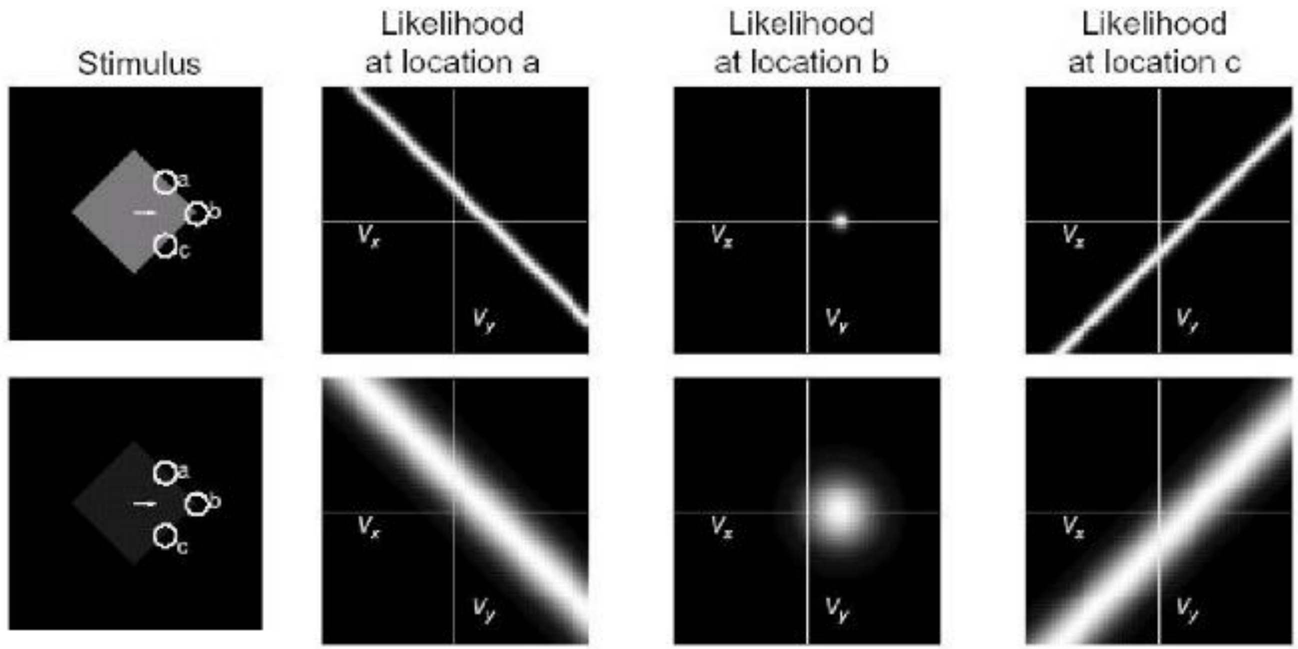
Lucas-Kanade with uncertainty:

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

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

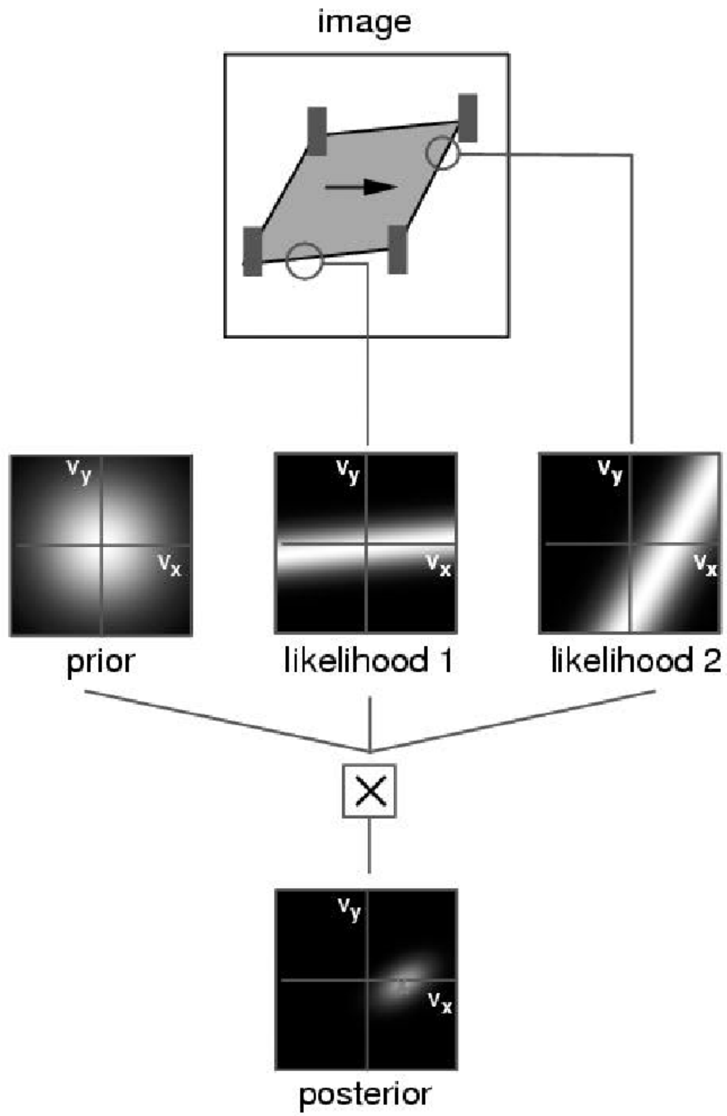


Figure 4: The response of the Bayesian estimator to a fat rhombus. (replotted from Weiss and Adelson 98)

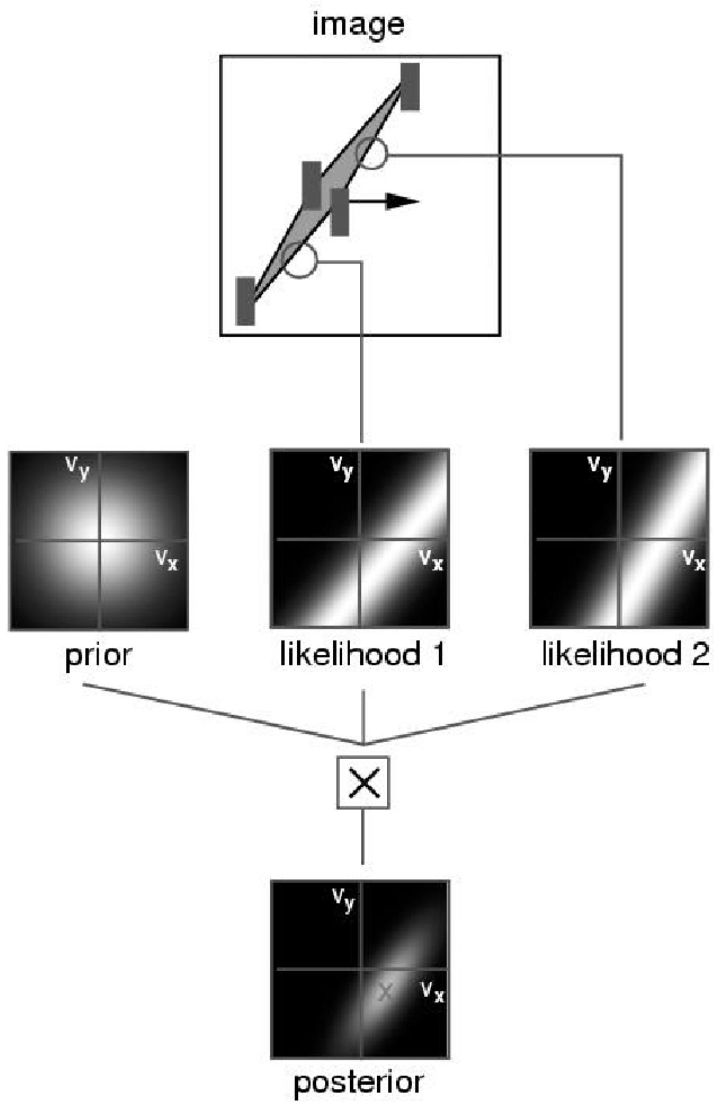
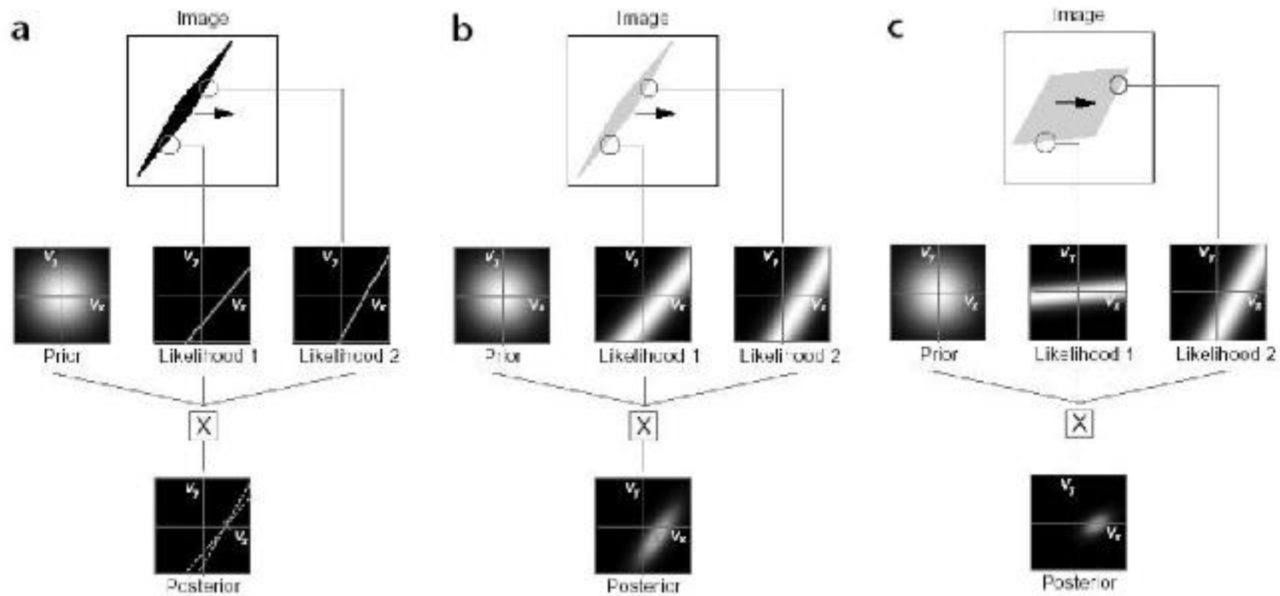


Figure 3: The response of the Bayesian estimator to a narrow rhombus. (replotted from Weiss and Adelson 98)

Effect of contrast



Thursday:

video compression