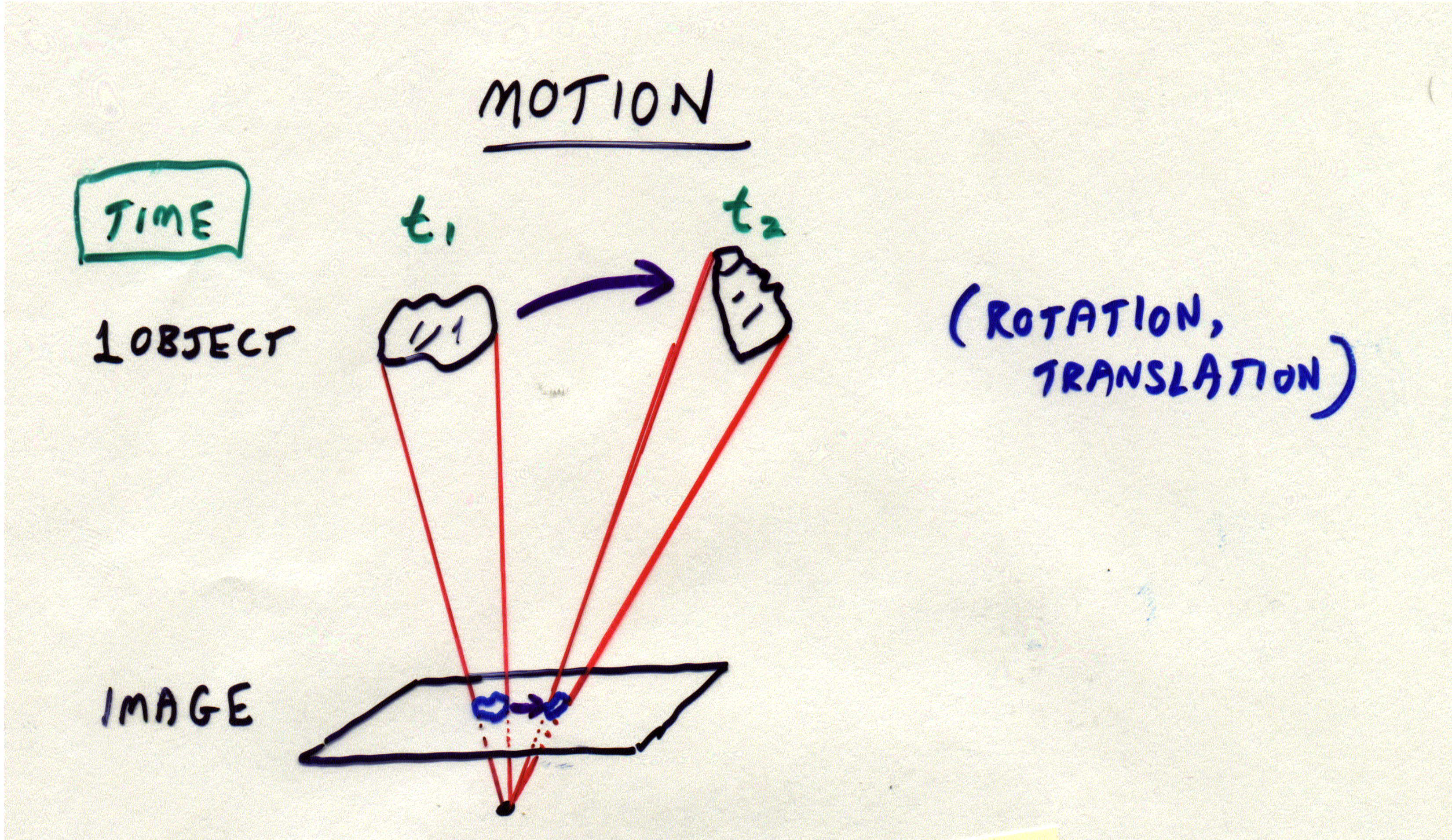
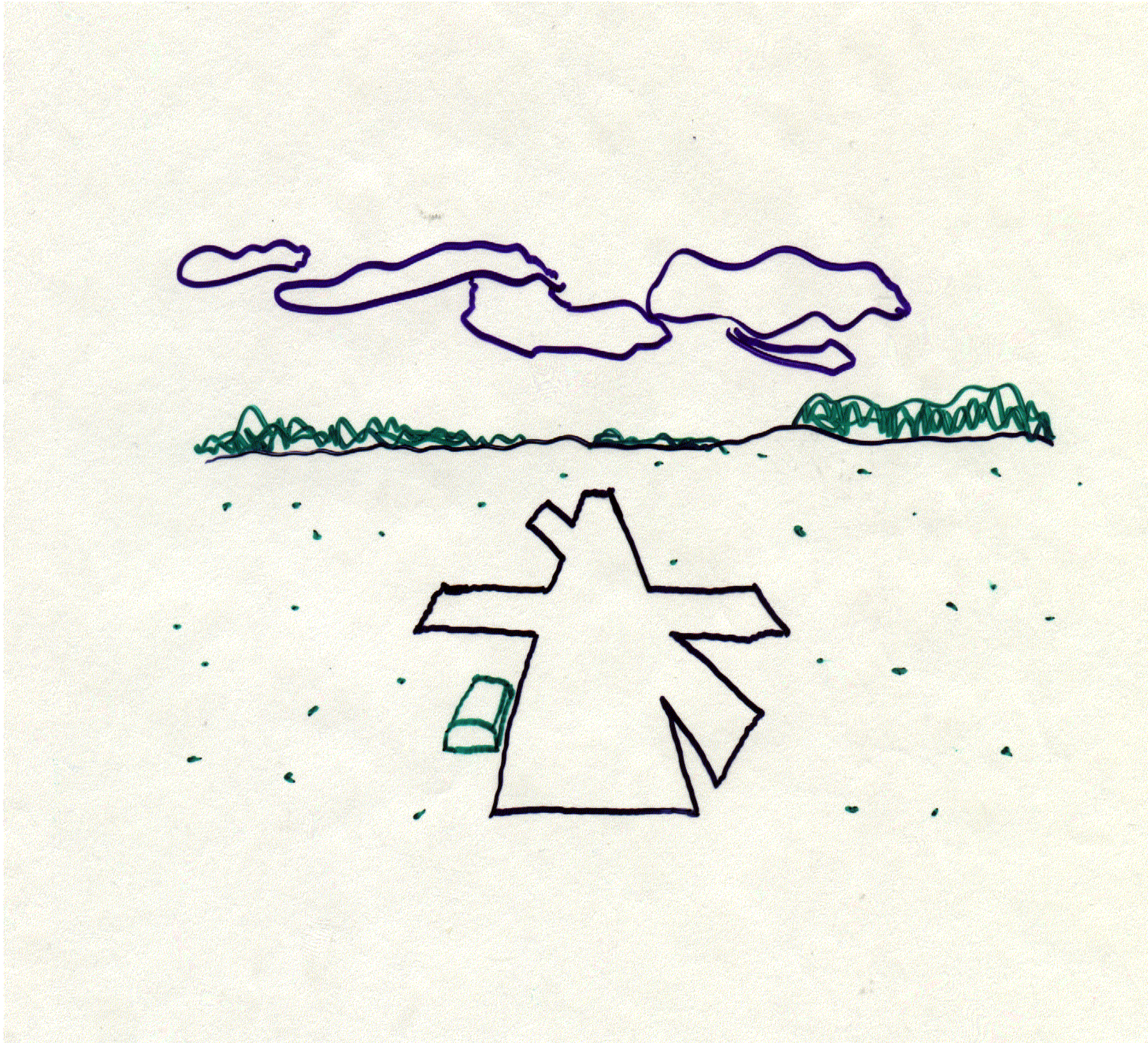
Motion analysis

Motion analysis: consider changes over time

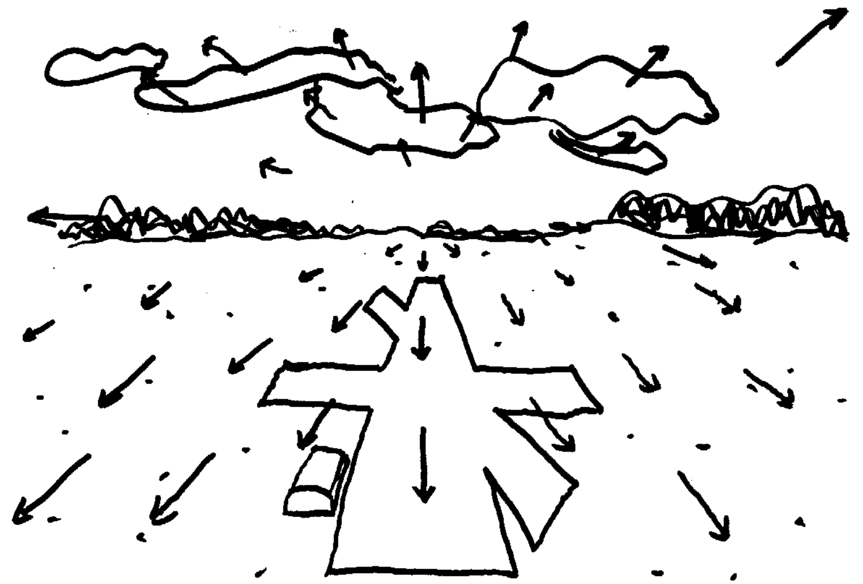
- Short time intervals
  - "Early warning system" based on sudden change
  - Image flow = instantaneous 2D velocity estimates (usually dense field)
  - Visual tracking to estimate new target locations in the image
  - Image segmentation to identify separately moving object surfaces
- Longer time intervals
  - 3D structure/shape from motion (SFM)
  - 3D velocity estimates (translational <u>and</u> rotational)
  - Estimate time to collision


MOTION
TIME
t_1
t_2
1 OBJECT
(ROTATION, TRANSLATION)
IMAGE



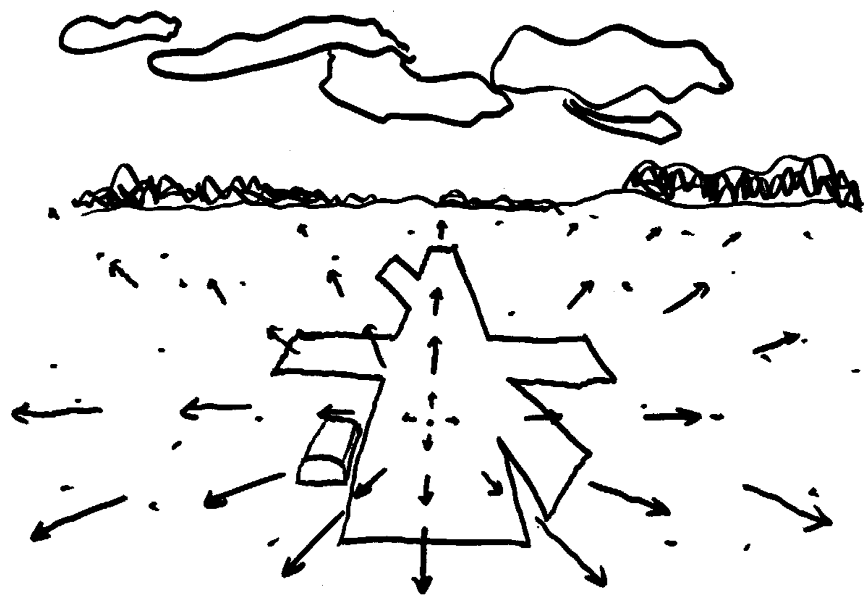
FIXATE HORIZON

OBSERVER MOVES FORWARD, PARALLEL TO GROUND



FOCUS OF EXPANSION
OR
(FOCUS OF CONTRACTION)

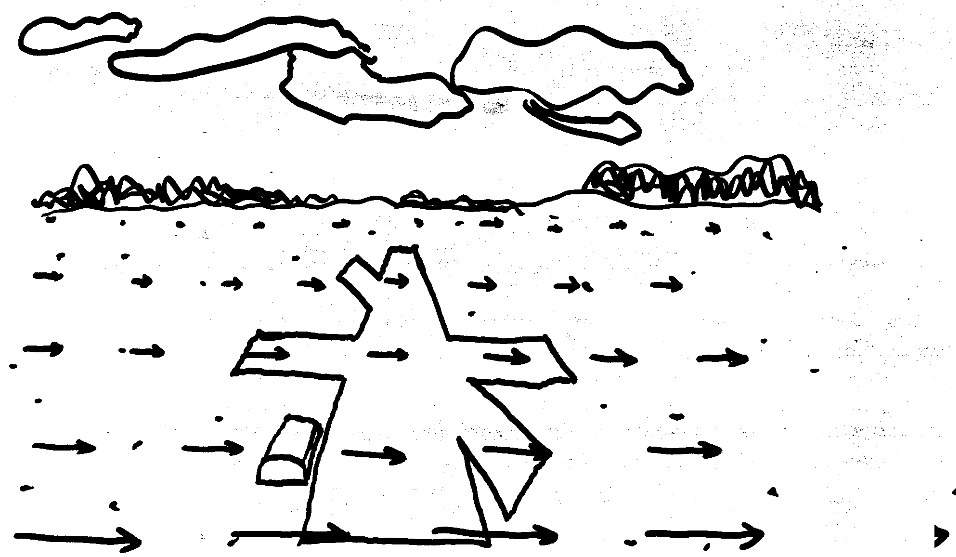
FIXATE A SPOT ON GROUND
OBSERVER MOVES TOWARD THAT POINT



- SURFACE NOT PARALLEL TO DIRECTION OF MOTION
  ⇒ ASYMMETRIC FLOW AROUND FOCUS

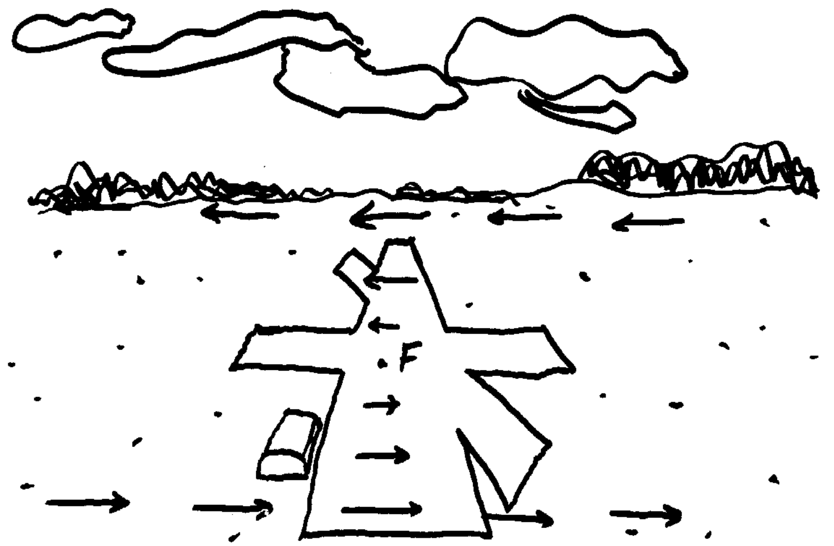
- NEARER SIDE HAS GREATER INCREASE IN MAGNITUDE

FIXATE HORIZON
OBSERVER MOVES TO LEFT



WHAT ABOUT CLOUDS?

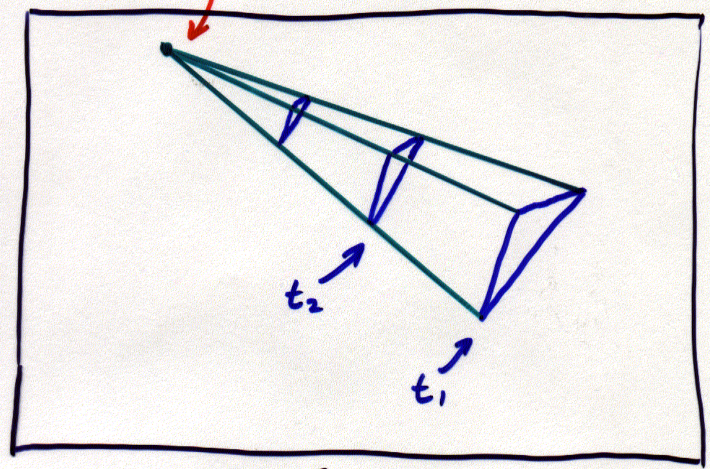
FIXATE A SPOT ON GROUND
OBSERVER MOVES TO THE LEFT



VELOCITY IN IMAGE IS ZERO AT F
DUE TO PURSUIT MOVEMENTS OF EYES.

FOCUS OF EXPANSION

or CONTRACTION



IMAGE

A RIGID OBJECT TRANSLATES WITH CONSTANT VELOCITY.

(COMPARE WITH THE VANISHING POINT CONCEPT.)

ONE SURFACE POINT AT 2 TIME INSTANTS

$$\bar{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$ = LOCATION OF SURFACE POINT AT TIME t_1

$$\bar{p}' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$ = LOCATION OF SAME SURFACE POINT AT TIME $t_2 > t_1$

$$\begin{bmatrix} X \\ Y \end{bmatrix}$$ = IMAGE OF POINT AT TIME t_1

$$\begin{bmatrix} X' \\ Y' \end{bmatrix}$$ = IMAGE OF POINT AT TIME t_2

PERSPECTIVE RELATION:

$$X = \frac{xf}{z}$$

$$Y = \frac{yf}{z}$$

$$X' = \frac{x'f}{z'}$$

$$Y' = \frac{y'f}{z'}$$

DEFINE $t_2 - t_1 = \Delta t$

FOCUS OF EXPANSION

- ASSUME OBJECT TRANSLATES SUCH THAT $\frac{dz}{dt} \neq 0$.
- ASSUME NO ROTATION.

$$\bar{p}' = \bar{p} + \left(\frac{d\bar{p}}{dt}\right)(\Delta t)$$

$$\Rightarrow \begin{bmatrix} \frac{X'z'}{f} \\ \frac{Y'z'}{f} \\ z' \end{bmatrix} = \begin{bmatrix} \frac{Xz}{f} \\ \frac{Yz}{f} \\ z \end{bmatrix} + \begin{bmatrix} \frac{dx}{dt}\Delta t \\ \frac{dy}{dt}\Delta t \\ \frac{dz}{dt}\Delta t \end{bmatrix}$$

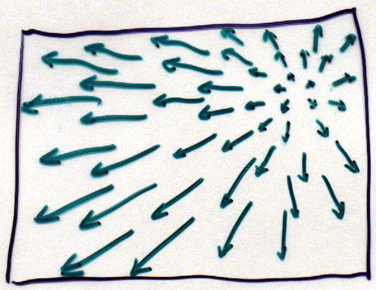
$$\Rightarrow \begin{cases} X' = \dfrac{\frac{Xz}{f} + \frac{dx}{dt}\Delta t}{z + \frac{dz}{dt}\Delta t} f \\[2ex] Y' = \dfrac{\frac{Yz}{f} + \frac{dy}{dt}\Delta t}{z + \frac{dz}{dt}\Delta t} f \end{cases}$$

$$\lim_{\Delta t \to \infty} X' = f \cdot \frac{\frac{dx}{dt}}{\frac{dz}{dt}} \approx f \frac{\frac{\Delta x}{\Delta t}}{\frac{\Delta z}{\Delta t}} = \boxed{f \frac{\Delta x}{\Delta z}}$$

$$\lim_{\Delta t \to \infty} Y' = f \cdot \frac{\frac{dy}{dt}}{\frac{dz}{dt}} \approx f \frac{\frac{\Delta y}{\Delta t}}{\frac{\Delta z}{\Delta t}} = \boxed{f \frac{\Delta y}{\Delta z}}$$

"EGOMOTION" $\triangleq$ SELF-MOTION

FOR AN AUTONOMOUS VEHICLE UNDERGOING TRANSLATIONAL MOTION, THE FOCUS OF EXPANSION CAN BE USED TO DETERMINE THE DIRECTION OF TRAVEL.



CAN TRY TO DETECT MOVING OBJECTS BY FINDING FLOW FIELDS THAT DIFFER FROM THIS SYMMETRIC PATTERN.

Image flow
(also called **optical flow** or **optic flow**)

- ***Goal:*** For each pixel in the image, determine a 2D velocity vector
  - Each velocity vector is assumed to represent the instantaneous 2D image velocity of a 3D moving object
  - This is like determining image disparities for 2 images taken at very short time intervals

Image flow recovery (Horn and Schunck, 1981)

- Consider the Taylor series expansion of *I:*

$$I\left(x+\Delta x, y+\Delta y, t+\Delta t\right)=I\left(x, y, t\right)+\frac{\partial I}{\partial x}\Delta x+\frac{\partial I}{\partial y}\Delta y+\frac{\partial I}{\partial t}\Delta t+h.o.t.$$

- Assume:
  - The movement is very small between successive images, so that we can ignore the higher-order terms
  - Locally, we can model the flow as translation only
- This leads to

$$\frac{\partial I}{\partial x}\Delta x+\frac{\partial I}{\partial y}\Delta y+\frac{\partial I}{\partial t}\Delta t=0$$

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t}+\frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t}+\frac{\partial I}{\partial t}=0 \longrightarrow \frac{\partial I}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial I}{\partial y}\frac{\partial y}{\partial t}+\frac{\partial I}{\partial t}=0$$

- Rewrite:

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

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

- This is the **image flow equation** (also called the "optical flow constraint equation")
- Now the goal is to determine the 2D velocity (u, v) at every pixel

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial t} \end{bmatrix}$$

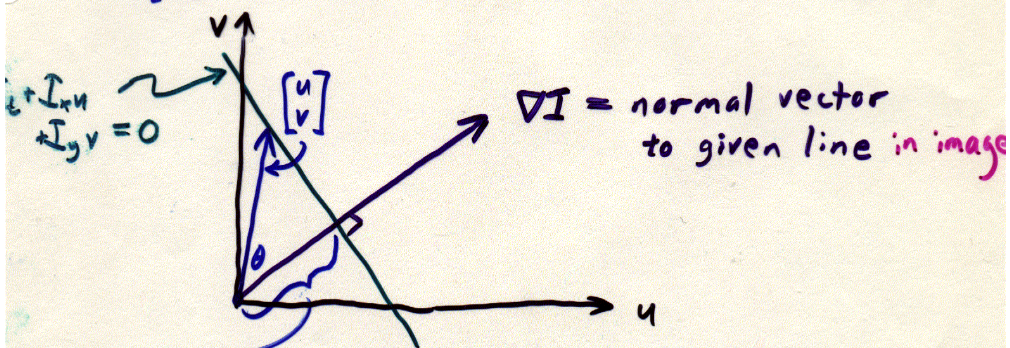
LET'S EXAMINE THIS CONSTRAINT:

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

$$\Rightarrow \quad \underbrace{-I_t}_{\text{TIME RATE OF CHANGE OF IMAGE INTENSITY}} = \underbrace{\nabla I}_{\text{IMAGE INTENSITY GRADIENT}} \cdot \underbrace{\begin{bmatrix} u \\ v \end{bmatrix}}_{\text{VELOCITY OF BRIGHTNESS VALUE IN IMAGE}}$$

THE CONSTRAINT DOES NOT DETERMINE $\begin{bmatrix} u \\ v \end{bmatrix}$ UNIQUELY!

$\begin{bmatrix} u \\ v \end{bmatrix}$ MUST LIE ON A LINE IN "VELOCITY SPACE":



WE CAN CALCULATE THE COMPONENT OF $\begin{bmatrix} u \\ v \end{bmatrix}$ IN DIRECTION OF ∇I:

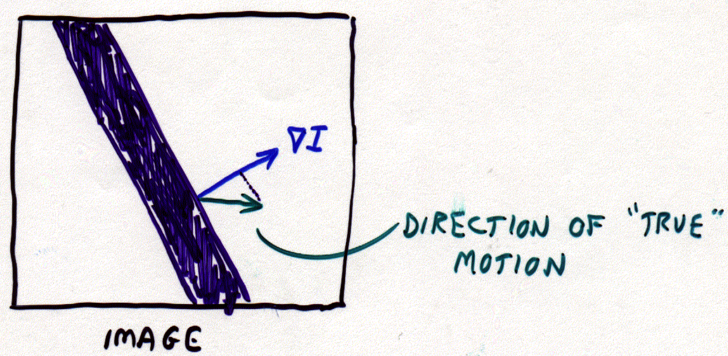
$$\left\| \begin{bmatrix} u \\ v \end{bmatrix} \right\| \cos\theta = \frac{-I_t}{\|\nabla I\|}$$

SO, FOR A POINT (x,y) IN THE IMAGE, WE CALCULATE THE MAGNITUDE OF THE OPTICAL FLOW VECTOR IN THE DIRECTION $\tan^{-1}\left(\frac{I_x}{I_y}\right)$ AS

$$\frac{-I_t}{\sqrt{I_x^2 + I_y^2}} .$$

UNFORTUNATELY, WE HAVE NO INFORMATION ABOUT THE VELOCITY ALONG THE ISOBRIGHTNESS CONTOUR IN $I(x,y)$. THIS IS CALLED THE $\boxed{\text{APERTURE}}$ $\boxed{\text{PROBLEM}}$

EXAMPLE:



ASSUME THAT THE DARK BAND MOVES HORIZONTALLY IN THE IMAGE. WE CAN ONLY CALCULATE THE SPEED OF THE BAND IN THE DIRECTION OF THE BRIGHTNESS GRADIENT ∇I!

CAUTION:

OPTICAL FLOW MAY NOT CORRESPOND TO THE 3D MOTION FIELD

EXAMPLE:

WHAT DO YOU SEE WHEN THIS OBJECT SPINS SLOWLY?

EXAMPLE:

"BARBER POLE"

EXAMPLE: LIGHTING CONDITIONS CHANGE, BUT OBJECT REMAINS STATIONARY

ALSO, NO OPTICAL FLOW INFORMATION IS AVAILABLE FOR FEATURELESS (NON-TEXTURED) SURFACES.

Balancing 2 constraints

- 1) How well the image flow equation holds:

$$e_c = \iint \left(I_x u + I_y v + I_t \right)^2 dx\,dy$$

- 2) Assume the flow field varies smoothly over *I*:

$$e_s = \iint \left[{u_x}^2 + {u_y}^2 + {v_x}^2 + {v_y}^2 \right] dx\,dy$$

or

$$e_s = \iint \left[\left(\nabla^2 u \right)^2 + \left(\nabla^2 v \right)^2 \right] dx\,dy$$

- Goal: find (u, v) throughout image to minimize

$$e = e_c + \lambda e_s$$

- Consider the integrand only:

$$e_{\text{int}} = \left(I_x u + I_y v + I_t\right)^2 + \lambda\left[\left(\nabla^2 u\right)^2 + \left(\nabla^2 v\right)^2\right]$$

- Approximate the Laplacian as $\nabla^2 u \approx u_{av} - u$

  (average of u-values of 4 neighbors)

- Now we have

$$e_{\text{int}} = \left(I_x u + I_y v + I_t\right)^2 + \lambda\left[\left(u - u_{av}\right)^2 + \left(v - v_{av}\right)^2\right]$$

- Revised goal: find (u, v) throughout image to minimize

$$e_{\text{int}} = \left(I_x u + I_y v + I_t\right)^2 + \lambda\left[\left(u - u_{av}\right)^2 + \left(v - v_{av}\right)^2\right]$$

- Strategy: solve for

$$\frac{\partial e_{\text{int}}}{\partial u} = 0 \qquad \frac{\partial e_{\text{int}}}{\partial v} = 0$$

$$\frac{\partial e_{\text{int}}}{\partial u} = 2\left(I_x u + I_y v + I_t\right) I_x + 2\lambda\left(u - u_{av}\right) = 0$$

$$\frac{\partial e_{\text{int}}}{\partial v} = 2\left(I_x u + I_y v + I_t\right) I_y + 2\lambda\left(v - v_{av}\right) = 0$$

- Collecting terms with respect to *u* and *v* yields

$$\left(I_x{}^2+\lambda\right)u+\left(I_xI_y\right)v+I_xI_t-\lambda u_{av}=0$$

$$\left(I_xI_y\right)u+\left(I_y{}^2+\lambda\right)v+I_yI_t-\lambda v_{av}=0$$

- Solve simultaneously for *u* and *v*:

$$u=u_{av}-I_x\left(\frac{I_xu_{av}+I_yv_{av}+I_t}{\lambda+I_x{}^2+I_y{}^2}\right)$$

$$v=v_{av}-I_y\left(\frac{I_xu_{av}+I_yv_{av}+I_t}{\lambda+I_x{}^2+I_y{}^2}\right)$$

Algorithm 1 (iterate on same 2 images):
(Horn and Schunck)

1. Initialize $u^{(0)} = v^{(0)} = 0$

   $k = 1$

2. Repeat until error criterion e is sufficiently small:

$$u^{(k)} := u_{av}{}^{(k-1)} - I_x \left(\frac{I_x u_{av}{}^{(k-1)} + I_y v_{av}{}^{(k-1)} + I_t}{\lambda + I_x{}^2 + I_y{}^2} \right)$$

$$v^{(k)} := v_{av}{}^{(k-1)} - I_y \left(\frac{I_x u_{av}{}^{(k-1)} + I_y v_{av}{}^{(k-1)} + I_t}{\lambda + I_x{}^2 + I_y{}^2} \right)$$

Algorithm 2 (iterate on successive image pairs):
(Horn and Schunck)

1. Initialize $u(x, y, 0) = v(x, y, 0) = 0$

   $t = 1$

2. Repeat until time t = latest-frame time:

$$u(x, y, t) := u_{av}(x, y, t-1) - I_x \left(\frac{I_x u_{av} + I_y v_{av} + I_t}{\lambda + I_x^{\,2} + I_y^{\,2}} \right)$$

$$v(x, y, t) := v_{av}(x, y, t-1) - I_y \left(\frac{I_x u_{av} + I_y v_{av} + I_t}{\lambda + I_x^{\,2} + I_y^{\,2}} \right)$$

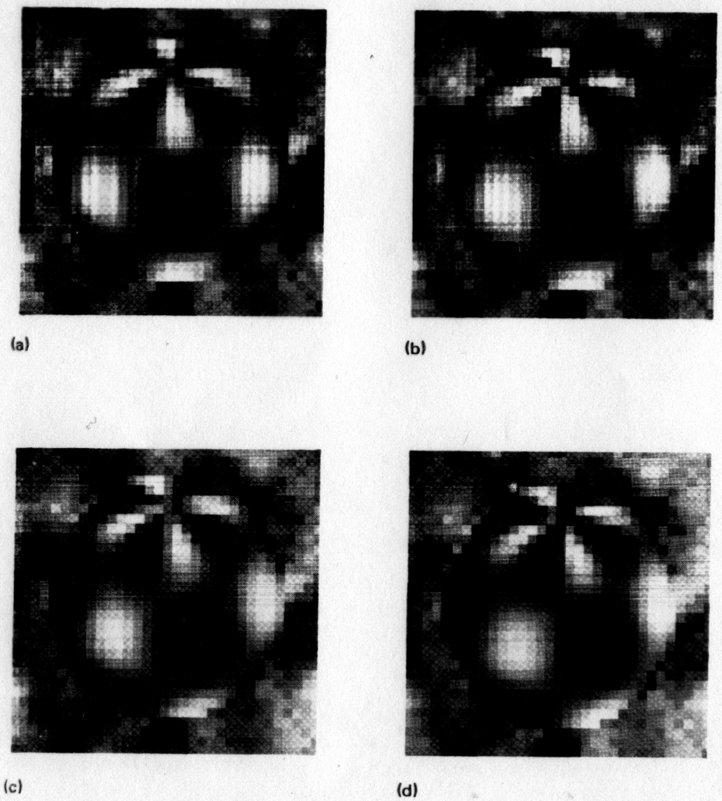


Figure 12-8. Four frames of a synthetic image sequence showing a sphere slowly rotating in front of a randomly patterned background.

12.7 Discontinuities in Optical Flow

There will be discontinuities in the optical flow on the silhouettes, where one object occludes another. We must detect these places if we are to prevent the method presented above from trying to continue the solution smoothly from one region to the other. This seems like a chicken-and-egg problem: If we have a good estimate of the optical flow, we can look for places where it changes very rapidly in order to segment the picture. On

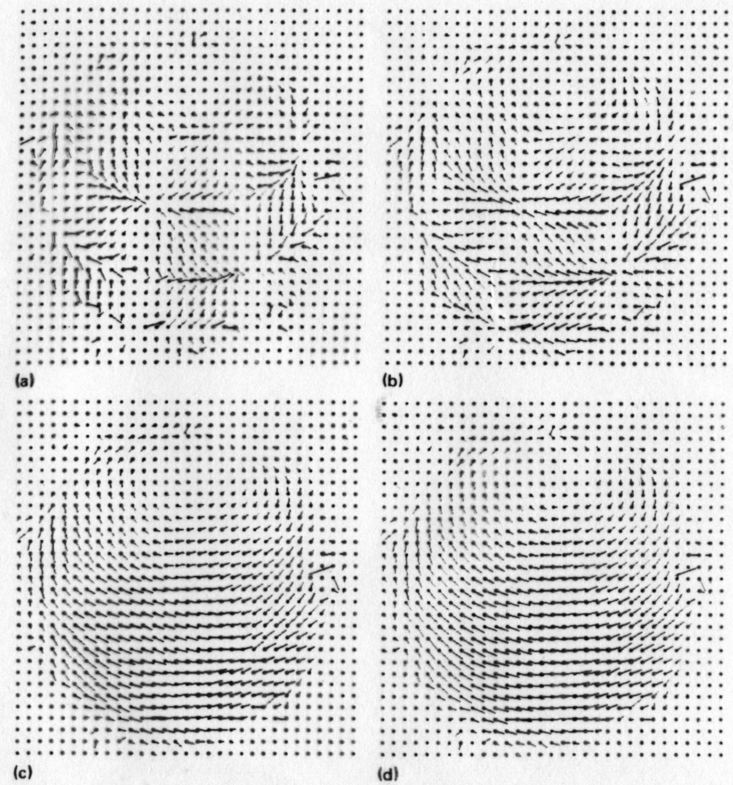


Figure 12-9. Estimates of the optical flow shown in the form of needle diagrams after 1, 4, 16, and 64 iterations of the algorithm.

the other hand, if we could segment the picture well, we would produce a better estimate of the optical flow. The solution to this dilemma is to incorporate the segmentation into the iterative solution for the optical flow. That is, after each iteration we look for places where the flow changes rapidly. At these places we set down marks that inhibit the next iteration from smoothly connecting the solution across the discontinuities. We first set the threshold for this decision very high in order to prevent premature carving up of the image. We reduce the threshold as better and better estimates of the optical flow become available.

We can get some feel for how well our assumption of smoothness holds

PROBLEMS WITH THIS METHOD

WHEN THE CONSTRAINTS ARE VIOLATED, THIS METHOD DOES NOT PERFORM WELL. THIS CAN OCCUR . . .

- AT OCCLUSION BOUNDARIES IN IMAGE
- AT SELF-OCCLUSION REGIONS
- WHEN MOTION CHARACTERISTICS CHANGE DRASTICALLY (e.g., COLLISIONS)

ALSO, INTENSITIES MAY NOT BE RELIABLE AS FEATURES:

- LACK OF IMAGE DETAIL
- CHANGES IN LIGHTING CONDITIONS
- SURFACE SPECULARITIES

Algorithm 3: (direct translational estimation)

- Recall the image flow equation: $\begin{bmatrix} I_x & I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = -I_t$

- Suppose we decide to constrain u and v to be constant over some small region of the image; then we can find a least-squares fit!

$$\begin{bmatrix} I_x(x_1, y_1) & I_y(x_1, y_1) \\ I_x(x_2, y_2) & I_y(x_2, y_2) \\ \vdots & \vdots \\ I_x(x_n, y_n) & I_y(x_n, y_n) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -I_t(x_1, y_1) \\ -I_t(x_2, y_2) \\ \vdots \\ -I_t(x_n, y_n) \end{bmatrix}$$

$$D \begin{bmatrix} u \\ v \end{bmatrix} = T \quad \longrightarrow \quad D^T D \begin{bmatrix} u \\ v \end{bmatrix} = D^T T \quad \longrightarrow \quad \begin{bmatrix} u \\ v \end{bmatrix} = \left(D^T D\right)^{-1} D^T T$$

Image flow: refinements

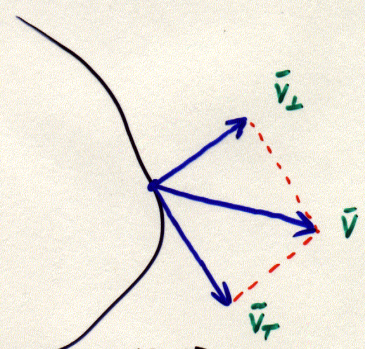
- Use the previous method with coarse-to-fine processing
  (This can handle large motion)

- Straight-forward search (usually area-based) to find best matches between successive images

AGAIN, CONSIDER THE
OPTICAL FLOW CONSTRAINT EQUATION:

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

FROM THIS, WE CAN ONLY DETERMINE THE COMPONENT OF VELOCITY THAT IS PARALLEL TO THE IMAGE GRADIENT.



$\bar{v} \triangleq \begin{bmatrix} u \\ v \end{bmatrix} \triangleq \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix}$ = ACTUAL VELOCITY IN IMAGE

$\bar{v}_\perp \triangleq$ VELOCITY COMPONENT PERPENDICULAR TO IMAGE CONTOUR

$\bar{v}_T \triangleq$ VELOCITY COMPONENT TANGENT TO CONTOUR

$$\bar{v} = \bar{v}_\perp + \bar{v}_T$$

OPTICAL FLOW FROM OBSERVED MOTION OF CONTOURS

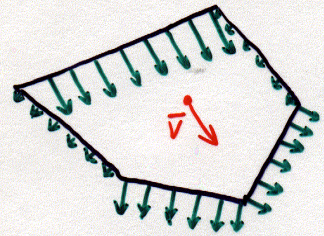
GIVEN: IMAGE CONTOURS
AND
PERPENDICULAR VELOCITY COMPONENTS $(\bar{V}_\perp)$

GOAL: DETERMINE ACTUAL VELOCITY FIELD $(\bar{V}_T)$
ALONG CONTOURS

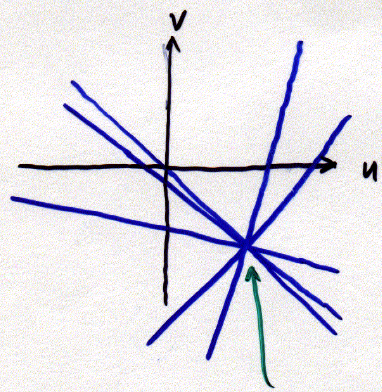
ASSUMPTIONS:

- THE IMAGE CONTOURS ARE THE PROJECTIONS OF 3D CURVES WHICH ARE IN MOTION.
- VELOCITY VECTORS ALONG IMAGE CONTOURS SHOULD BE CONSISTENT WITH THE UNDERLYING, DENSE VELOCITY FIELD.

EXAMPLE 1: RIGID TRANSLATION IN IMAGE PLANE



← MOVING IMAGE REGION

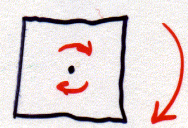


← VELOCITY SPACE

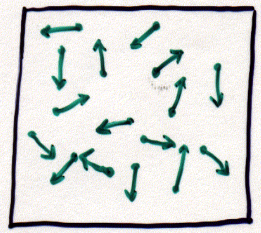
THE CONSTRAINT LINES INTERSECT AT A SINGLE POINT IN VELOCITY SPACE, INDICATING THE ACTUAL REGION VELOCITY $\bar{v}$.

CAUTION:

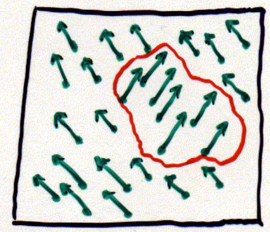
THIS WILL NOT USUALLY WORK FOR <u>ROTATION</u>



2D MOTION DISCRIMINATION



SEGMENTATION
NOT POSSIBLE



SEGMENTATION
POSSIBLE

- OPTICAL FLOW DIRECTION IS IMPORTANT FOR OBJECT DETECTION:

  DIRECTION BOUNDARIES (2D) $\Rightarrow$ OBJECT BOUNDARIES (3D)

- USUALLY, SEGMENTATION BASED ON FLOW DIRECTION SEPARATES SURFACES THAT MOVE INDEPENDENTLY.