Motion and optical flow

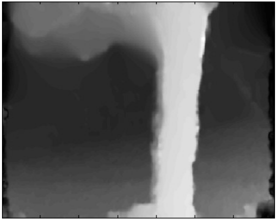
DD2423 Image Analysis and Computer Vision

Mårten Björkman

Division of Robotics, Perception and Learning
School of Electrical Engineering and Computer Science

December 13, 2021

What about motion?



- Measuring sideways motion is very much like stereo.
- A single camera at two different instances in time can be seen as two cameras at two different locations.

Motion is more complex

However

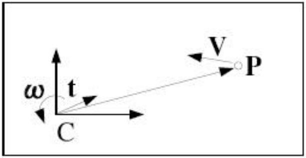
- Motion can be in any direction, not just along “epipolar lines”.
- One cannot tell how large the image motion is. For disparities one can have an idea of maximum and minimum values.
- The image motion arises from both
  - the motion of the camera (ego-motion), and
  - the motion of things in the scene (independent motion).

- Consider an observer moving with an angular velocity ω and translational velocity T in a static environment.
- In relation to the observer, a 3D point $P = (X, Y, Z)^\top$ moves as

$$\dot{P} = -T - \omega \times P \qquad (1)$$

or explicitly

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{pmatrix} = -\begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} - \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$



Motion field due to ego-motion

- The projection in the image is

$$\left\{\begin{array}{l} x = f\frac{X}{Z} \\ y = f\frac{Y}{Z} \end{array}\right. \Rightarrow \left\{\begin{array}{l} \dot{x} = f\frac{Z\dot{X}-X\dot{Z}}{Z^2} \\ \dot{y} = f\frac{Z\dot{Y}-Y\dot{Z}}{Z^2} \end{array}\right.$$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \frac{f}{Z}\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} - \frac{f}{Z}\begin{pmatrix} X \\ Y \end{pmatrix}\frac{\dot{Z}}{Z} \qquad (2)$$

Motion field due to ego-motion

- Combine the two equations (1) and (2)

$$\begin{cases} \dot{X} = -(T_x + \omega_y Z - \omega_z Y) \\ \dot{Y} = -(T_y + \omega_z X - \omega_x Z) \\ \dot{Z} = -(T_z + \omega_x Y - \omega_y X) \end{cases}$$

$$\begin{aligned} \frac{f}{Z}\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} &= -\frac{f}{Z}\begin{pmatrix} T_x + \omega_y Z - \omega_z Y \\ T_y + \omega_z X - \omega_x Z \end{pmatrix} \\ &= -\frac{f}{Z}\begin{pmatrix} T_x \\ T_y \end{pmatrix} - \begin{pmatrix} \omega_y f - \omega_z y \\ -\omega_x f + \omega_z x \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \frac{f}{Z}\begin{pmatrix} X \\ Y \end{pmatrix}\frac{\dot{Z}}{Z} &= -\begin{pmatrix} x \\ y \end{pmatrix}\frac{1}{Z}(T_z + \omega_x Y - \omega_y X) \\ &= -\begin{pmatrix} x \\ y \end{pmatrix}\left(\frac{T_z}{Z} + \omega_x \frac{y}{f} - \omega_y \frac{x}{f}\right) \\ &= -\begin{pmatrix} x \\ y \end{pmatrix}\left(\frac{T_z}{Z}\right) - \begin{pmatrix} x \\ y \end{pmatrix}\left(\omega_x \frac{y}{f} - \omega_y \frac{x}{f}\right) \end{aligned}$$

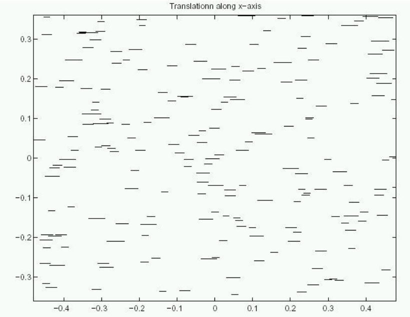
Motion field due to ego-motion

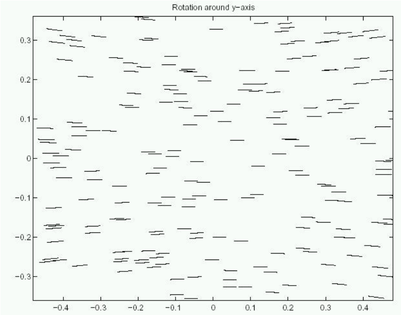
Add these together $\Rightarrow$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \underbrace{\frac{f}{Z}\begin{pmatrix} -T_x + \frac{x}{f}T_z \\ -T_y + \frac{y}{f}T_z \end{pmatrix}}_{\text{translation, scaled by 1/Z}} + \underbrace{\begin{pmatrix} \omega_x \frac{xy}{f} - \omega_y(f + \frac{x^2}{f}) + \omega_z y \\ \omega_x(f + \frac{y^2}{f}) - \omega_y \frac{xy}{f} + \omega_z x \end{pmatrix}}_{\text{rotation, independent of depth}}$$

- Translational component depends inversely on depth, scaling ambiguity: T and Z can be recovered only up to a scale.
- Rotational component does not depend on depth – impossible to estimate depth without translation.

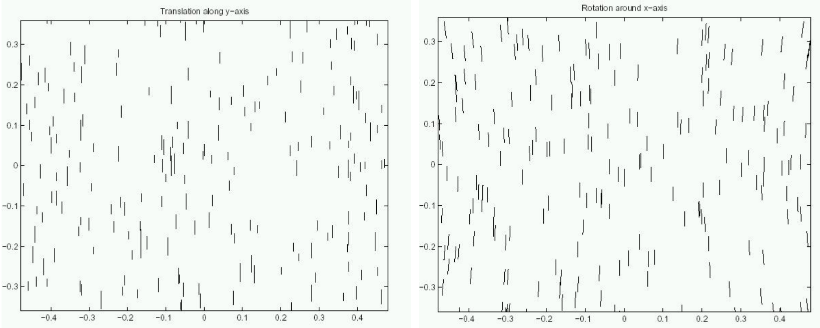
Motion flows





Translation T_x, Rotation ω_y

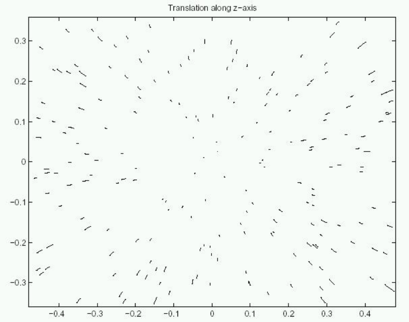
Translational and rotational flows are very similar.



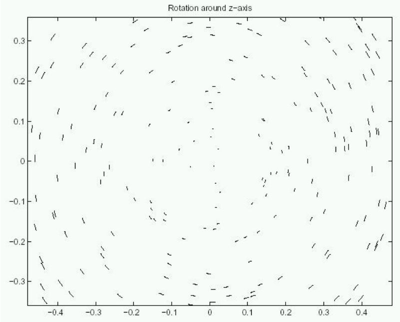
Translation T_y, Rotation ω_x

Translational and rotational flows are very similar.

Motion flows



Translation T_z,



Rotation ω_z

Except for fowards motion and rotation around optical axis.

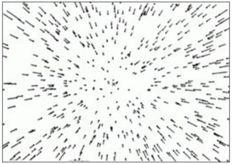
Translational motion

$$\frac{1}{f}\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = -\frac{1}{Z}\begin{pmatrix} T_x \\ T_y \end{pmatrix} + \begin{pmatrix} x/f \\ y/f \end{pmatrix}\frac{T_z}{Z} = \frac{1}{Z}\begin{pmatrix} x/f \cdot T_z - T_x \\ y/f \cdot T_z - T_y \end{pmatrix}$$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} x_{FOE} \\ y_{FOE} \end{pmatrix} = \frac{f}{T_z}\begin{pmatrix} T_x \\ T_y \end{pmatrix}$$

- The flow-field expands from a point, the Focus of Expansion.

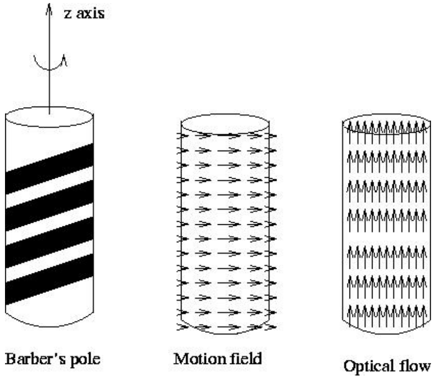
$$\begin{pmatrix} x_{FOE} \\ y_{FOE} \\ f \end{pmatrix} = \frac{f}{T_z}\begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}$$



- Conclusion: Translation direction can be seen directly in image.
- Comparison: Flow vectors in 3D are parallel and parallel lines "intersect" in a Vanishing Point.

Optical flow

- Optical flow is the apparent motion of brightness patterns.
- Generally, optical flow corresponds to motion field, but not always.
- For example, motion field and optical flow of a rotating barber's pole are different, as illustrated in the figure



Optical flow constraint equation

- Denote the intensity of a single scene point by $I(x(t), y(t), t)$.
- This is a function of three variables, as we now have spatio-temporal variation in our signal.
- To see how I changes, we differentiate with respect to time t:

$$\frac{dI}{dt} = I_x \frac{dx}{dt} + I_y \frac{dy}{dt} + I_t$$

Optical flow constraint equation

- If we assume that the image intensity of each visible scene point is constant over time (brightness constancy), we have

$$\frac{dI}{dt} = 0$$

which implies

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

where the partial derivatives of I are denoted by subscripts, and u and v are the x and y components of the optical flow vector.

- This equation is called the *optical flow constraint equation*, since it expresses a constraint on the components the optical flow.

Optical flow constraint equation

- The optical flow constraint equation can be rewritten as

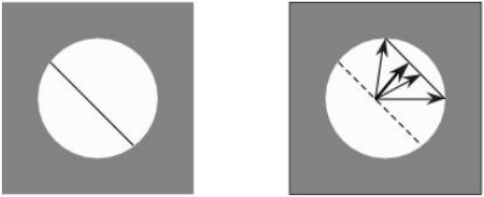
$$(I_x, I_y) \cdot (u, v) = -I_t$$

- Thus, the component of the image velocity in the direction of the image intensity gradient at the image of a scene point is

$$(u, v) = \frac{-I_t}{I_x^2 + I_y^2}(I_x, I_y)$$

- We cannot determine the component of the optical flow along an edge. This ambiguity is known as the *aperture problem*.

The aperture problem



(a) (b)

(a) A line feature observed through a small aperture at time t.

(b) At $t + \delta t$ the line has moved. It is not possible to determine exactly where, since along the line everything looks the same.

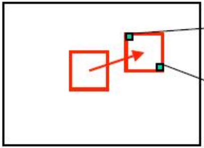
- Normal flow: component of flow perpendicular to the line.

Local smoothness: Lucas, Kanade (1984)

- One pixel is not enough (one equation, two unknowns).

$$(I_x, I_y) \cdot (u, v) = -I_t$$

- Assume local smoothness (constancy) in a windows.



$$\begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \vdots & \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_{t1} \\ I_{t2} \\ \vdots \end{bmatrix}$$

$$\boxed{A\vec{u} = b}$$

Local smoothness: Lucas, Kanade (1984)

- Goal: Minimize $||Au - b||^2$

$$f(u) = (Au - b)^T (Au - b) = u^T A^T Au - 2u^T A^T b - b^T b$$

$$f'(u) = 2A^T Au - 2A^T b = 0$$

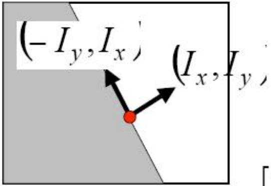
$$u = (A^T A)^{-1} A^T b$$

$$A^T A = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix}$$

- $A^T A$ is the Second moment matrix used for corner detection.
- We need this matrix to be invertible $\Rightarrow$ No zero eigenvalues.

Behaviour due to Second moment matrix

- Edge ➔ $A^T A$ becomes singular



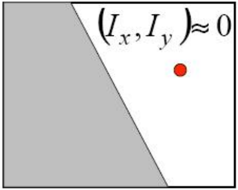
$$\begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} \begin{bmatrix} -I_y \\ I_x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

⇓

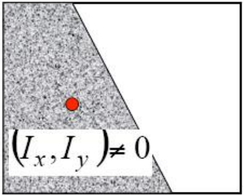
$\begin{bmatrix} -I_y \\ I_x \end{bmatrix}$ is eigenvector with eigenvalue 0

Behaviour due to Second moment matrix

- Homogeneous ➔ $A^T A \approx 0$ ➔ 0 eigenvalues



- Textured regions ➔ two high eigenvalues



Affine motion

- Instead of assuming that motion is constant within local window, assume that it satisfies an affine model

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

- Substitute it in the optical flow constraint equation

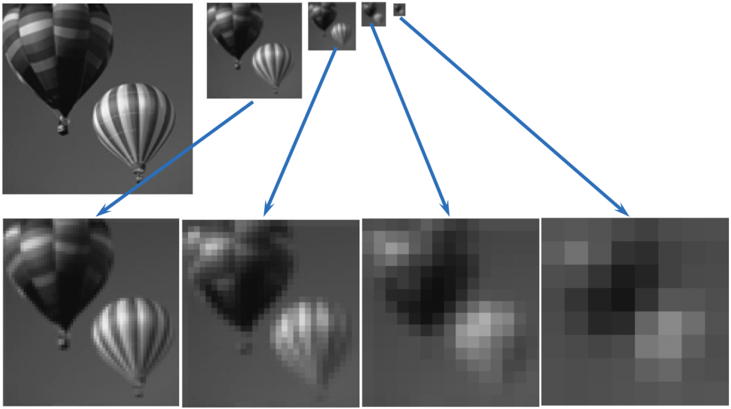
$$I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t = 0$$

- Apply least square minimization over a local window to find the unknown 6 parameters.

Some problems remain

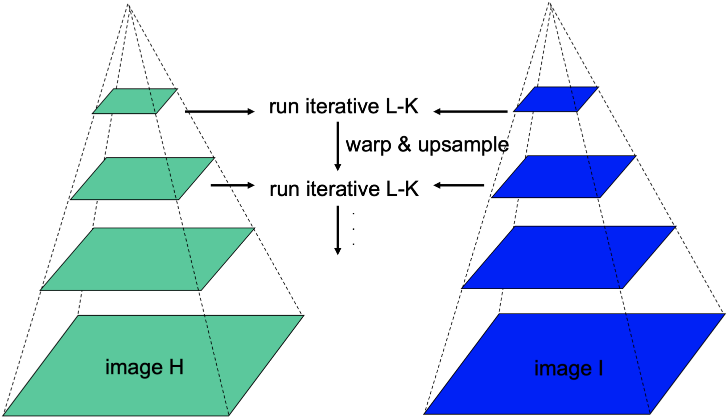
- Lucas & Kanade (and similar) methods only handle flow smaller than the standard deviation of the Gaussian blurring filter.
- Possible solutions:
  - Iteratively shift windows for matching over time.
  - Search coarse-to-fine using Gaussian pyramids.

Coarse-to-fine Lucas & Kanade



- Create pyramid by successively blurring and subsampling image.

Coarse-to-fine Lucas & Kanade

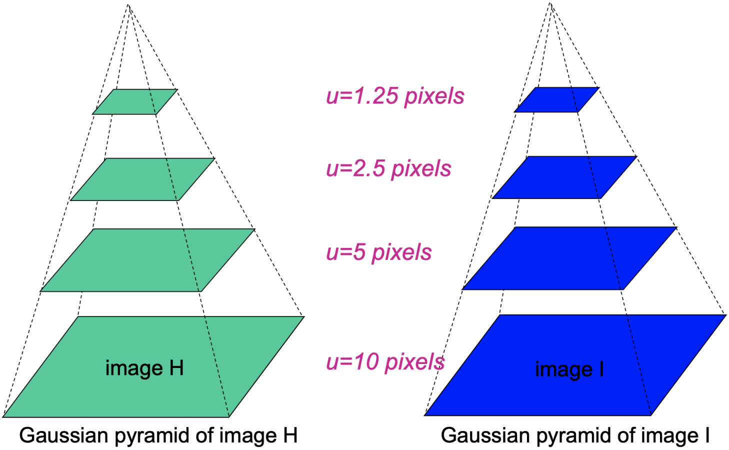


Gaussian pyramid of image H

Gaussian pyramid of image I

- Run Lucas & Kanade iteratively and upsample estimated optical flow from coarse to finer scale.

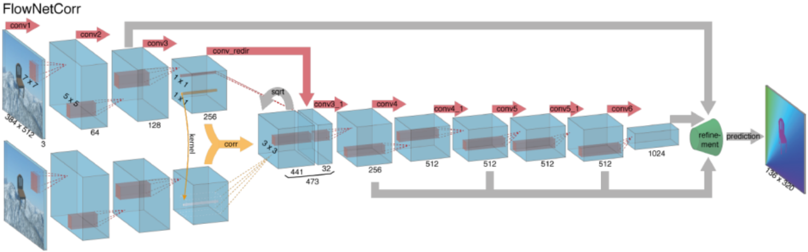
Coarse-to-fine Lucas & Kanade



- Gradually the maximum amount of optical flow can increase and capture more realistic image motion.

FlowNet: optical flow with CNNs (2015)

This is one out of many possible deep network based solutions.



- Two images as input applied to a conventional CNN.
- CNN features are then correlated, followed by a sequence of unpooling and convolution layers.
- Final optical flow predictions are based on outputs from multiple layers, similar to coarse-to-fine Lucas & Kanade.

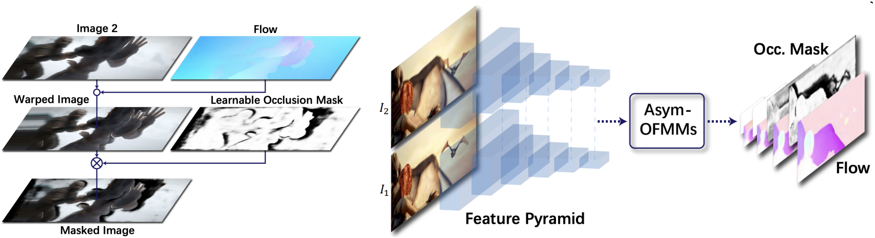
Dosovitskiy et al., “FlowNet: Learning Optical Flow with Convolutional Networks”, ICCV 2015.

FlowNet example

What is most challenging is to find sharp boundaries and still get accurate optical flow estimates. It's very hard to get both.

MaskFlowNet (2020)

Currently one of the best solutions.



- First extract multi-scale features data in a feature pyramid.
- Then estimates optical flow AND occlusion mask at each scale.
- Propagates flow coarse-to-fine similarly to Lucas & Kanade.

Zhao at al., ”MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask”, CVPR 2020.

At the end .....

- Only one exercise session left on Wednesday.
- Please fill in the course evaluation that will appear on KTH Social!

What about the exam?

- Still planned to be held on campus!
- 13 January, 08:00-13:00
- Three kinds of questions
  - C: Concept questions
  - P1: Easier problem questions
  - P2: More complex questions
- Exam registration is needed. If not registered, visit service center.
- Allowed tools: calculator and mathematical handbook (e.g Beta)

Labs and exercises

Laboratory exercises:

- Reread what you did. What were you supposed to have learned?
- It will help you on both practical and theoretical parts of the exam.

Exercise sessions:

- Go through the problems!
- Likely that something similar is on the exam.

AND, please fill in “Course evaluation form”!

Summary of good questions

- Why is motion more complex than stereo?
- What is a Focus of Expansion?
- What is Motion field and what is Optical flow?
- How do you derive the optical flow constraint?
- What is a Second moment matrix?

Readings

- Szeliski: Chapters 9.1 and 9.3