

RWTH AACHEN  
UNIVERSITY

# Computer Vision - Lecture 20

## Motion and Optical Flow

28.01.2016

Computer Vision WS 15/16

Bastian Leibe  
RWTH Aachen  
<http://www.vision.rwth-aachen.de>  
leibe@vision.rwth-aachen.de

Many slides adapted from K. Grauman, S. Seitz, R. Szeliski, M. Pollefeys, S. Lazebnik

RWTH AACHEN  
UNIVERSITY

## Course Outline

- Image Processing Basics
- Segmentation & Grouping
- Object Recognition
- Local Features & Matching
- Object Categorization
- 3D Reconstruction
  - Epipolar Geometry and Stereo Basics
  - Camera calibration & Uncalibrated Reconstruction
  - Active Stereo
- Motion
  - Motion and Optical Flow
- 3D Reconstruction (Reprise)
  - Structure-from-Motion

2

RWTH AACHEN  
UNIVERSITY

## Recap: Epipolar Geometry - Calibrated Case

$x \cdot [t \times (Rx')] = 0 \Rightarrow x^T E x' = 0$  with  $E = [t_x] R$

↓

**Essential Matrix**  
(Longuet-Higgins, 1981)

Slide credit: Svetlana Lazebnik

RWTH AACHEN  
UNIVERSITY

## Recap: Epipolar Geometry - Uncalibrated Case

$\hat{x}^T E \hat{x}' = 0 \Rightarrow x^T F x' = 0$  with  $F = K^{-T} E K'^{-1}$

$x = K \hat{x}$   
 $x' = K' \hat{x}'$

↓

**Fundamental Matrix**  
(Faugeras and Luong, 1992)

Slide credit: Svetlana Lazebnik

RWTH AACHEN  
UNIVERSITY

## Recap: The Eight-Point Algorithm

$x = (u, v, 1)^T, x' = (u', v', 1)^T$

$(u, v, 1) \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \Rightarrow [u'u, u'v, u'u', uv', v'v', v'u, v, u, v, 1] \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$

↓

$\begin{bmatrix} u_1'u_1 & u_1'v_1 & u_1'u' & u_1v_1 & u_1v_1' & u_1v_1 & u_1 & v_1 & 1 \\ u_2'u_1 & u_2'v_1 & u_2'u' & u_2v_1 & u_2v_1' & u_2v_1 & u_2 & v_2 & 1 \\ u_3'u_1 & u_3'v_1 & u_3'u' & u_3v_1 & u_3v_1' & u_3v_1 & u_3 & v_3 & 1 \\ u_4'u_1 & u_4'v_1 & u_4'u' & u_4v_1 & u_4v_1' & u_4v_1 & u_4 & v_4 & 1 \\ u_5'u_1 & u_5'v_1 & u_5'u' & u_5v_1 & u_5v_1' & u_5v_1 & u_5 & v_5 & 1 \\ u_6'u_1 & u_6'v_1 & u_6'u' & u_6v_1 & u_6v_1' & u_6v_1 & u_6 & v_6 & 1 \\ u_7'u_1 & u_7'v_1 & u_7'u' & u_7v_1 & u_7v_1' & u_7v_1 & u_7 & v_7 & 1 \\ u_8'u_1 & u_8'v_1 & u_8'u' & u_8v_1 & u_8v_1' & u_8v_1 & u_8 & v_8 & 1 \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

↓

Solve using... SVD!

This minimizes:  $Af = 0 \Rightarrow \sum_{i=1}^N (x_i^T F x'_i)^2$

Slide adapted from Svetlana Lazebnik

RWTH AACHEN  
UNIVERSITY

## Recap: Normalized Eight-Point Algorithm

1. Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels.
2. Use the eight-point algorithm to compute  $F$  from the normalized points.
3. Enforce the rank-2 constraint using SVD.
 

Set  $d_{33}$  to zero and reconstruct  $F'$

$$F = UDV^T = U \begin{bmatrix} d_{11} & & & \\ & d_{22} & & \\ & & d_{33} & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} v_{11} & \dots & v_{13} \\ \vdots & \ddots & \vdots \\ v_{31} & \dots & v_{33} \end{bmatrix}^T$$
4. Transform fundamental matrix back to original units: if  $T$  and  $T'$  are the normalizing transformations in the two images, then the fundamental matrix in original coordinates is  $T^T F T'$ .

Slide credit: Svetlana Lazebnik

RWTH AACHEN UNIVERSITY

## Practical Considerations

Small Baseline

Large Baseline

1. Role of the baseline
  - > Small baseline: large depth error
  - > Large baseline: difficult search problem
- Solution
  - > Track features between frames until baseline is sufficient.

Computer Vision WS 15/16 | Slide adapted from Steve Seitz | B. Leibe | 15

RWTH AACHEN UNIVERSITY

## Topics of This Lecture

- Introduction to Motion
  - > Applications, uses
- Motion Field
  - > Derivation
- Optical Flow
  - > Brightness constancy constraint
  - > Aperture problem
  - > Lucas-Kanade flow
  - > Iterative refinement
  - > Global parametric motion
  - > Coarse-to-fine estimation
  - > Motion segmentation
- KLT Feature Tracking

Computer Vision WS 15/16 | B. Leibe | 16

RWTH AACHEN UNIVERSITY

## Video

- A video is a sequence of frames captured over time
- Now our image data is a function of space (x, y) and time (t)

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe | 17

RWTH AACHEN UNIVERSITY

## Motion and Perceptual Organization

- Sometimes, motion is the only cue...

<ul style="list-style-type: none"> <li> Not grouped</li> <li> Proximity</li> <li> Similarity</li> <li> Similarity</li> <li> Common Fate</li> <li> Common Region</li> </ul>	<ul style="list-style-type: none"> <li> Parallelism</li> <li> Symmetry</li> <li> Continuity</li> <li> Closure</li> </ul>
--	--

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe | 18

RWTH AACHEN UNIVERSITY

## Motion and Perceptual Organization

- Sometimes, motion is foremost cue

Computer Vision WS 15/16 | Slide credit: Kristen Grauman | B. Leibe | 19

RWTH AACHEN UNIVERSITY

## Motion and Perceptual Organization


- Even "impoverished" motion data can evoke a strong percept

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe | 20

RWTH AACHEN UNIVERSITY

## Motion and Perceptual Organization

- Even “impovertished” motion data can evoke a strong percept



21

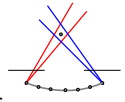
Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik B. Leibe

RWTH AACHEN UNIVERSITY

## Uses of Motion

- Estimating 3D structure
  - Directly from optic flow
  - Indirectly to create correspondences for SfM
- Segmenting objects based on motion cues
- Learning dynamical models
- Recognizing events and activities
- Improving video quality (motion stabilization)



22

Computer Vision WS 15/16

Slide adapted from Svetlana Lazebnik B. Leibe

RWTH AACHEN UNIVERSITY

## Motion Estimation Techniques

- Direct methods
  - Directly recover image motion at each pixel from spatio-temporal image brightness variations
  - Dense motion fields, but sensitive to appearance variations
  - Suitable for video and when image motion is small
- Feature-based methods
  - Extract visual features (corners, textured areas) and track them over multiple frames
  - Sparse motion fields, but more robust tracking
  - Suitable when image motion is large (10s of pixels)

23

Computer Vision WS 15/16

Slide credit: Steve Seitz B. Leibe

RWTH AACHEN UNIVERSITY

## Topics of This Lecture

- Introduction to Motion
  - Applications, uses
- Motion Field
  - Derivation
- Optical Flow
  - Brightness constancy constraint
  - Aperture problem
  - Lucas-Kanade flow
  - Iterative refinement
  - Global parametric motion
  - Coarse-to-fine estimation
  - Motion segmentation
- KLT Feature Tracking

24

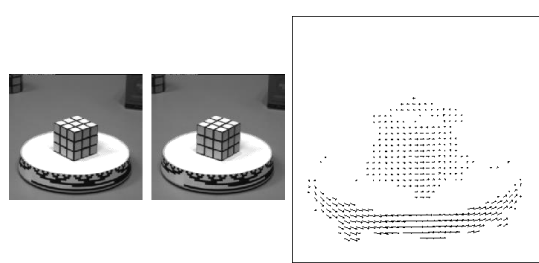
Computer Vision WS 15/16

B. Leibe

RWTH AACHEN UNIVERSITY

## Motion Field

- The motion field is the projection of the 3D scene motion into the image



25

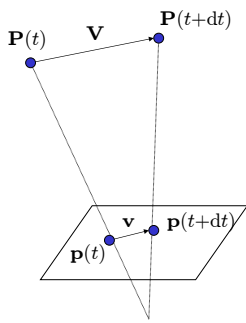
Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik B. Leibe

RWTH AACHEN UNIVERSITY

## Motion Field and Parallax

- $\mathbf{P}(t)$  is a moving 3D point
- Velocity of 3D scene point:  $\mathbf{V} = d\mathbf{P}/dt$
- $\mathbf{p}(t) = (x(t), y(t))$  is the projection of  $\mathbf{P}$  in the image.
- Apparent velocity  $\mathbf{v}$  in the image: given by components  $v_x = dx/dt$  and  $v_y = dy/dt$
- These components are known as the *motion field* of the image.



26

Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik B. Leibe

Quotient rule:  
 $(f/g)' = (g f' - f g')/g^2$

### Motion Field and Parallax

$\mathbf{V} = [V_x, V_y, V_z]$     $\mathbf{p} = f \frac{\mathbf{P}}{Z}$

To find image velocity  $\mathbf{v}$ , differentiate  $\mathbf{p}$  with respect to  $t$  (using quotient rule):

$$\mathbf{v} = f \frac{Z\mathbf{V} - V_z\mathbf{P}}{Z^2} = \frac{f\mathbf{V} - V_z\mathbf{p}}{Z}$$

$$v_x = \frac{fV_x - V_zx}{Z} \quad v_y = \frac{fV_y - V_zy}{Z}$$

- Image motion is a function of both the 3D motion ( $\mathbf{V}$ ) and the depth of the 3D point ( $Z$ ).

27

RWTH AACHEN UNIVERSITY

### Motion Field and Parallax

- Pure translation:  $\mathbf{V}$  is constant everywhere

$$v_x = \frac{fV_x - V_zx}{Z} \quad \mathbf{v} = \frac{1}{Z}(\mathbf{v}_0 - V_z\mathbf{p}),$$

$$v_y = \frac{fV_y - V_zy}{Z} \quad \mathbf{v}_0 = (fV_x, fV_y)$$

28

RWTH AACHEN UNIVERSITY

### Motion Field and Parallax

- Pure translation:  $\mathbf{V}$  is constant everywhere

$$\mathbf{v} = \frac{1}{Z}(\mathbf{v}_0 - V_z\mathbf{p}),$$

$$\mathbf{v}_0 = (fV_x, fV_y)$$

- $V_z$  is nonzero:
  - Every motion vector points toward (or away from)  $\mathbf{v}_0$ , the vanishing point of the translation direction.

29

RWTH AACHEN UNIVERSITY

### Motion Field and Parallax

- Pure translation:  $\mathbf{V}$  is constant everywhere

$$\mathbf{v} = \frac{1}{Z}(\mathbf{v}_0 - V_z\mathbf{p}),$$

$$\mathbf{v}_0 = (fV_x, fV_y)$$

- $V_z$  is nonzero:
  - Every motion vector points toward (or away from)  $\mathbf{v}_0$ , the vanishing point of the translation direction.
- $V_z$  is zero:
  - Motion is parallel to the image plane, all the motion vectors are parallel.
- The length of the motion vectors is inversely proportional to the depth  $Z$ .

30

RWTH AACHEN UNIVERSITY

### Topics of This Lecture

- Introduction to Motion
  - Applications, uses
- Motion Field
  - Derivation
- Optical Flow
  - Brightness constancy constraint
  - Aperture problem
  - Lucas-Kanade flow
  - Iterative refinement
  - Global parametric motion
  - Coarse-to-fine estimation
  - Motion segmentation
- KLT Feature Tracking

31

RWTH AACHEN UNIVERSITY

### Optical Flow

- Definition: optical flow is the *apparent* motion of brightness patterns in the image.
- Ideally, optical flow would be the same as the motion field.
- Have to be careful: apparent motion can be caused by lighting changes without any actual motion.
  - Think of a uniform rotating sphere under fixed lighting vs. a stationary sphere under moving illumination.

32

RWTH AACHEN UNIVERSITY

## Apparent Motion ≠ Motion Field

(a) (b)

**Figure 12-2.** The optical flow is not always equal to the motion field. In (a) a smooth sphere is rotating under constant illumination—the image does not change, yet the motion field is nonzero. In (b) a fixed sphere is illuminated by a moving source—the shading in the image changes, yet the motion field is zero.

33

Computer Vision WS 15/16 | Slide credit: Kristen Grauman | B. Leibe | Figure from Horn book

RWTH AACHEN UNIVERSITY

## Estimating Optical Flow

$I(x,y,t-1)$        $I(x,y,t)$

- Given two subsequent frames, estimate the apparent motion field  $u(x,y)$  and  $v(x,y)$  between them.
- Key assumptions
  - Brightness constancy:** projection of the same point looks the same in every frame.
  - Small motion:** points do not move very far.
  - Spatial coherence:** points move like their neighbors.

34

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe

RWTH AACHEN UNIVERSITY

## The Brightness Constancy Constraint

$(x,y)$  displacement =  $(u,v)$

$I(x,y,t-1)$        $I(x,y,t)$

- Brightness Constancy Equation:**

$$I(x, y, t-1) = I(x+u(x, y), y+v(x, y), t)$$
- Linearizing the right hand side using Taylor expansion:
 
$$I(x, y, t-1) \approx I(x, y, t) + I_x \cdot u(x, y) + I_y \cdot v(x, y)$$
- Hence,  $I_x \cdot u + I_y \cdot v + I_t \approx 0$ 
  - Spatial derivatives:**  $I_x, I_y$
  - Temporal derivative:**  $I_t$

35

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe

RWTH AACHEN UNIVERSITY

## The Brightness Constancy Constraint

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

- How many equations and unknowns per pixel?
  - One equation, two unknowns
- Intuitively, what does this constraint mean?
 
$$\nabla I \cdot (u, v) + I_t = 0$$
- The component of the flow perpendicular to the gradient (i.e., parallel to the edge) is unknown

If  $(u, v)$  satisfies the equation, so does  $(u+u', v+v')$  if  $\nabla I \cdot (u', v') = 0$

36

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe

RWTH AACHEN UNIVERSITY

## The Aperture Problem

Perceived motion

37

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe

RWTH AACHEN UNIVERSITY

## The Aperture Problem

Actual motion

38

Computer Vision WS 15/16 | Slide credit: Svetlana Lazebnik | B. Leibe



RWTH AACHEN UNIVERSITY

## Eigenvectors of $A^T A$

$$A^T A = \begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} = \sum \begin{bmatrix} I_x \\ I_y \end{bmatrix} \begin{bmatrix} I_x & I_y \end{bmatrix} = \sum \nabla I (\nabla I)^T$$

- Haven't we seen an equation like this before?
- Recall the Harris corner detector:  $M = A^T A$  is the second moment matrix.
- The eigenvectors and eigenvalues of  $M$  relate to edge direction and magnitude.
  - The eigenvector associated with the larger eigenvalue points in the direction of fastest intensity change.
  - The other eigenvector is orthogonal to it.

45

RWTH AACHEN UNIVERSITY

## Interpreting the Eigenvalues

- Classification of image points using eigenvalues of the second moment matrix:

46

RWTH AACHEN UNIVERSITY

## Edge

$$\sum \nabla I (\nabla I)^T$$

- Gradients very large or very small
- Large  $\lambda_1$ , small  $\lambda_2$

47

RWTH AACHEN UNIVERSITY

## Low-Texture Region

$$\sum \nabla I (\nabla I)^T$$

- Gradients have small magnitude
- Small  $\lambda_1$ , small  $\lambda_2$

48

RWTH AACHEN UNIVERSITY

## High-Texture Region

$$\sum \nabla I (\nabla I)^T$$

- Gradients are different, large magnitude
- Large  $\lambda_1$ , large  $\lambda_2$

49

RWTH AACHEN UNIVERSITY

## Per-Pixel Estimation Procedure

- Let  $M = \sum (\nabla I)(\nabla I)^T$  and  $b = \begin{bmatrix} -\sum I_x I_x \\ -\sum I_x I_y \end{bmatrix}$
- 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
  - Trivially singular if the summation is over a single pixel or if there is no texture
  - I.e., only normal flow is available (aperture problem)
- Corners and textured areas are OK

50

RWTH AACHEN UNIVERSITY

## Iterative Refinement

1. Estimate velocity at each pixel using one iteration of Lucas and Kanade estimation.
 
$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

$$A^T A \qquad \qquad \qquad A^T b$$
2. Warp one image toward the other using the estimated flow field.
  - (Easier said than done)
3. Refine estimate by repeating the process.

51

RWTH AACHEN UNIVERSITY

## Optical Flow: Iterative Refinement

Initial guess:  $d_0 = 0$   
Estimate:  $d_1 = d_0 + \tilde{d}$

(using  $d$  for displacement here instead of  $u$ )

52

RWTH AACHEN UNIVERSITY

## Optical Flow: Iterative Refinement

Initial guess:  $d_1$   
Estimate:  $d_2 = d_1 + \tilde{d}$

(using  $d$  for displacement here instead of  $u$ )

53

RWTH AACHEN UNIVERSITY

## Optical Flow: Iterative Refinement

Initial guess:  $d_2$   
Estimate:  $d_3 = d_2 + \tilde{d}$

(using  $d$  for displacement here instead of  $u$ )

54

RWTH AACHEN UNIVERSITY

## Optical Flow: Iterative Refinement

$f_1(x - d_3) \approx f_2(x)$

(using  $d$  for displacement here instead of  $u$ )

55

RWTH AACHEN UNIVERSITY

## Optical Flow: Iterative Refinement

- Some Implementation Issues:
  - Warping is not easy (ensure that errors in warping are smaller than the estimate refinement).
  - Warp one image, take derivatives of the other so you don't need to re-compute the gradient after each iteration.
  - Often useful to low-pass filter the images before motion estimation (for better derivative estimation, and linear approximations to image intensity).

56



RWTH AACHEN UNIVERSITY

## Extension: Global Parametric Motion Models

Computer Vision WS 15/16

Slide credit: Steve Seitz

57

RWTH AACHEN UNIVERSITY

## Example: Affine Motion

$$u(x, y) = a_1 + a_2x + a_3y$$

$$v(x, y) = a_4 + a_5x + a_6y$$

- Substituting into the brightness constancy equation:

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

Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik

58

RWTH AACHEN UNIVERSITY

## Example: Affine Motion

$$u(x, y) = a_1 + a_2x + a_3y$$

$$v(x, y) = a_4 + a_5x + a_6y$$

- Substituting into the brightness constancy equation:

$$I_x(a_1 + a_2x + a_3y) + I_y(a_4 + a_5x + a_6y) + I_t \approx 0$$

- Each pixel provides 1 linear constraint in 6 unknowns.
- Least squares minimization:

$$Err(\vec{a}) = \sum [I_x(a_1 + a_2x + a_3y) + I_y(a_4 + a_5x + a_6y) + I_t]^2$$

Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik

59

RWTH AACHEN UNIVERSITY

## Problem Cases in Lucas-Kanade

- The motion is large (larger than a pixel)
  - Iterative refinement, coarse-to-fine estimation
- A point does not move like its neighbors
  - Motion segmentation
- Brightness constancy does not hold
  - Do exhaustive neighborhood search with normalized correlation.

Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik

60

RWTH AACHEN UNIVERSITY

## Dealing with Large Motions

Computer Vision WS 15/16

Slide credit: Svetlana Lazebnik

61

RWTH AACHEN UNIVERSITY

## Temporal Aliasing

- Temporal aliasing causes ambiguities in optical flow because images can have many pixels with the same intensity.
- I.e., how do we know which 'correspondence' is correct?

- To overcome aliasing: coarse-to-fine estimation.

Computer Vision WS 15/16

Slide credit: Steve Seitz

62

Computer Vision WS 15/16

**Idea: Reduce the Resolution!**

Slide credit: Svetlana Lazebnik

B. Leibe

63

Computer Vision WS 15/16

**Coarse-to-fine Optical Flow Estimation**

$u=1.25$  pixels

$u=2.5$  pixels

$u=5$  pixels

$u=10$  pixels

Image 1

Image 2

Gaussian pyramid of image 1

Gaussian pyramid of image 2

Slide credit: Steve Seitz

B. Leibe

64

Computer Vision WS 15/16

**Coarse-to-fine Optical Flow Estimation**

Run iterative L-K

Warp & upsample

Run iterative L-K

Image 1

Image 2

Gaussian pyramid of image 1

Gaussian pyramid of image 2

Slide credit: Steve Seitz

B. Leibe

65

Computer Vision WS 15/16

**Dense Optical Flow**

Dense measurements can be obtained by adding smoothness constraints.

Color map

(c) Thomas Brox 2009

T. Brox, C. Bregler, J. Malik, [Large displacement optical flow](#), CVPR'09, Miami, USA, June 2009.

Slide credit: Steve Seitz

B. Leibe

83

Computer Vision WS 15/16

**Summary**

- Motion field: 3D motions projected to 2D images; dependency on depth.
- Solving for motion with
  - Sparse feature matches
  - Dense optical flow
- Optical flow
  - Brightness constancy assumption
  - Aperture problem
  - Solution with spatial coherence assumption

Slide credit: Kristen Grauman

B. Leibe

84

Computer Vision WS 15/16

**References and Further Reading**

- Here is the original paper by Lucas & Kanade
  - B. Lucas and T. Kanade. [An iterative image registration technique with an application to stereo vision](#). In *Proc. IJCAI*, pp. 674-679, 1981.
- And the original paper by Shi & Tomasi
  - J. Shi and C. Tomasi. [Good Features to Track](#). CVPR 1994.

Slide credit: Steve Seitz

B. Leibe

85