Lecture 24:
Image Sequence Analysis I:
Local Methods

Contents

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Introduction

Basic Problem

- given: image sequence $f(x, y, z)$,
  where (x, y) specifies the location and z denotes time
- wanted: displacement vector field of the image structures:
  optic flow (optischer Fluss) $\begin{pmatrix} u(x,y,z) \\ v(x,y,z) \end{pmatrix}$

Such correspondence problems are key issues in computer vision.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Similar Correspondence Problems:

- computing the displacements *(disparities)* between the two images of a stereo pair
- matching *(registration)* of medical images that are obtained with different modalities, parameter settings or at different times

What is Optic Flow Good for?

- estimation of motion parameters in robotics
- reconstruction of the 3-D world from an image sequence *(structure-from-motion)*
- tracking of moving objects, e.g. human body motion
- efficient video coding

M I A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

A Frequent Assumption: Grey Value Constancy

Corresponding image structures have the same grey value:
The optic flow between frame z and $z+1$ satisfies

$$f(x+u,\ y+v,\ z+1) \;=\; f(x,y,z)$$

Linearisation by a Taylor Expansion:

- If (u,v) is small and f varies slowly, we have

$$\begin{aligned} 0 &= f(x+u,y+v,z+1) - f(x,y,z) \\ &\approx f_x(x,y,z)\,u + f_y(x,y,z)\,v + f_z(x,y,z) \end{aligned}$$

  where subscripts denote partial derivatives.
- This yields the *linearised optic flow constraint (OFC)*

$$f_x u + f_y v + f_z = 0$$

Introduction (4)

How Realisitic are These Assumptions?

- The grey value constancy assumption is often surprisingly realistic:
  Usually, illumination changes take place very slowly, i.e. over many frames.
  There are also more complicated models that take into account illumination changes.

- The linearisation assumption is violated more frequently:
  Conventional video cameras often suffer from temporal undersampling, i.e. they produce displacements over several pixels. Remedies:
  - use original OFC without linearisation (models become more difficult)
  - spatial downsampling (after lowpass filtering!)

- Another practical problem: *interlacing*.
  Camera records odd and even rows at different times. This may create artifacts for moving objects. Remedies:
  - consider only one of the two subimages
  - buy more expensive cameras without interlacing *(progressive scan cameras)*

Introduction (5)

The Aperture Problem (Blendenproblem, Aperturproblem)

- OFC specifies only the flow component $w_n = (u,v)\frac{\nabla f}{|\nabla f|}$ parallel to the spatial gradient $\nabla f := (f_x, f_y)^\top$ (and thus orthogonal to the image edges):

  $$0 \;=\; f_x u + f_y v + f_z \;=\; (u,v)\nabla f + f_z \;=\; w_n |\nabla f| + f_z$$

  Therefore, this so-called *normal flow* has the magnitude

  $$w_n = -\frac{f_z}{|\nabla f|}$$

- Nonuniqueness: One can add arbitrary flow components orthogonal to ∇f without violating the OFC.

$\Longrightarrow$ **additional assumptions necessary**

Specifying different additional constraints leads to different methods.

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

(a) Top left: Image from the Hamburg taxi sequence. **(b) Top right:** Normal flow magnitude without presmoothing the derivatives. **(c) Bottom left:** Presmoothing with a Gaussian with standard deviation $\sigma = 2$. **(d) Bottom right:** $\sigma = 4$. Author: J. Weickert (2001).

M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

The Spatial Approach of Lucas and Kanade (1981)

- Additional assumption for dealing with the aperture problem:
  The optic flow in (x_0, y_0) at time z_0 can be approximated by a *constant* vector (u, v) within some neighbourhood $B_\rho(x_0, y_0)$ of radius ρ.

- Desired flow minimises the *local* energy

$$E(u,v) = \frac{1}{2} \int_{B_\rho(x_0,y_0)} (f_x u + f_y v + f_z)^2 \, dx \, dy$$

- Computing the partial derivatives with respect to u and v gives

$$0 \stackrel{!}{=} \frac{\partial E}{\partial u} = \int_{B_\rho} f_x (f_x u + f_y v + f_z) \, dx \, dy$$

$$0 \stackrel{!}{=} \frac{\partial E}{\partial v} = \int_{B_\rho} f_y (f_x u + f_y v + f_z) \, dx \, dy$$

M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

The Spatial Approach of Lucas and Kanade (2)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

- u and v are constants, that can be moved out of the integral.
  This yields the linear system

$$\begin{pmatrix} \int\limits_{B_\rho} f_x^2\,dx\,dy & \int\limits_{B_\rho} f_x f_y\,dx\,dy \\ \int\limits_{B_\rho} f_x f_y\,dx\,dy & \int\limits_{B_\rho} f_y^2\,dx\,dy \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\int\limits_{B_\rho} f_x f_z\,dx\,dy \\ -\int\limits_{B_\rho} f_y f_z\,dx\,dy \end{pmatrix}$$

- Often one replaces the box filter with a "hard" window $B_\rho(x,y)$ by a "smooth" convolution with a Gaussian K_ρ:

$$\begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * (f_x f_z) \\ -K_\rho * (f_y f_z) \end{pmatrix}$$

- Thus, the Lucas–Kanade method solves a 2×2 linear system of equations.
  The (spatial) structure tensor J_ρ serves as system matrix (cf. Lecture 19).

The Spatial Approach of Lucas and Kanade (3)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

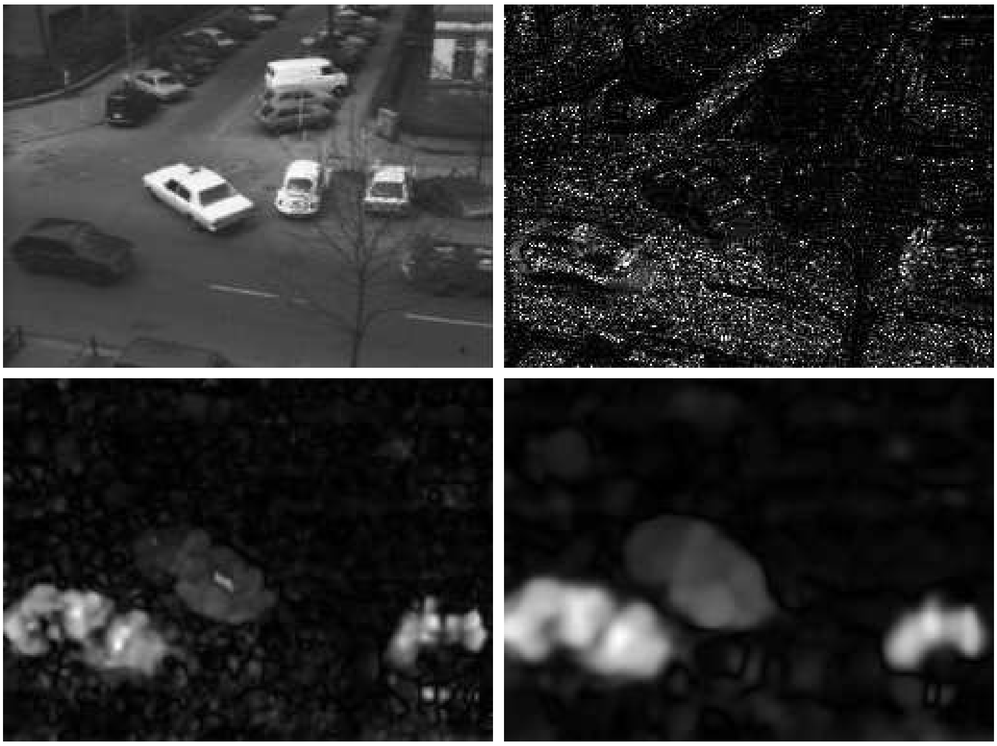
When Does This System Have No Unique Solution ?

- $\text{rank}(J_\rho) = 0$ (two vanishing eigenvalues):
  Happens if the spatial gradient vanishes in the entire neighbourhood.
  Simple criterion: $\text{tr}\, J_\rho \;=\; j_{11} + j_{22} \;\le\; \varepsilon.$
  Nothing can be said in this case.

- $\text{rank}(J_\rho) = 1$ (one vanishing eigenvalue):
  Happens if we have the same (nonvanishing) gradient within the entire neighbourhood.
  Then then both equations are linearly dependent (infinitely many solutions).
  Simple criterion: $\det J_\rho \;=\; j_{11} j_{22} - j_{12}^2 \;\le\; \varepsilon$ (while $\text{tr}\, J_\rho > \varepsilon$).
  In this case the aperture problem persists. One can only compute the normal flow:

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = w_n \cdot \frac{\nabla f}{|\nabla f|} = \frac{-1}{f_x^2 + f_y^2} \begin{pmatrix} f_x f_z \\ f_y f_z \end{pmatrix}$$

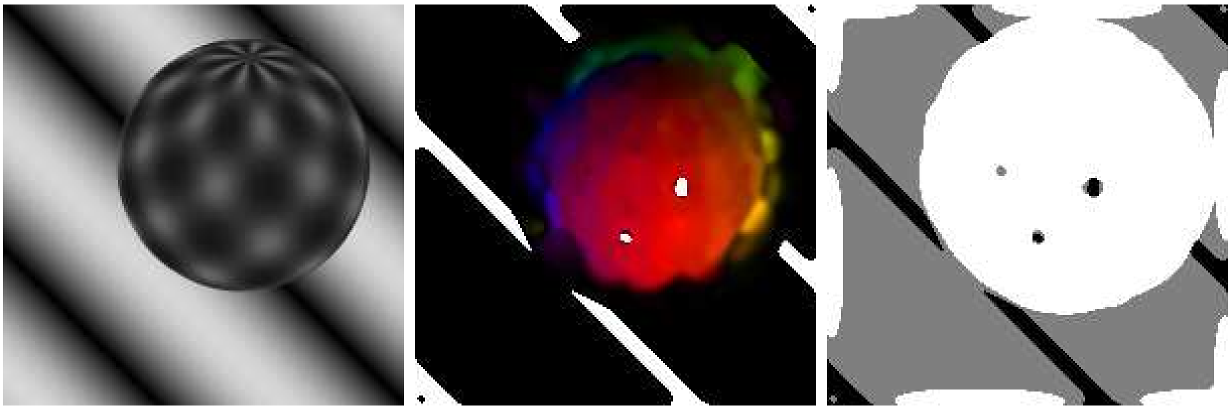
The Spatial Approach of Lucas and Kanade (4)



(a) Top left: Image from the Hamburg taxi sequence. **(b) Top right:** Normal flow magnitude. **(c) Bottom left:** Optic flow magnitude using the Lucas-Kanade method with $\rho = 2$. **(d) Bottom right:** Same with $\rho = 4$. Author: J. Weickert (2001).

The Spatial Approach of Lucas and Kanade (5)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26



(a) Left: Image from a synthetic sequence: The sphere rotates in front of a static background. **(b) Middle:** False colour representation of the optic flow using the Lucas–Kanade method. **(c) Right:** Flow classification: black=no information (gradient too small), grey=aperture problem (gradient too uniform), white=full flow (space-variant gradient). Author: J. Weickert (2001).

Advantages

- simple and fast method
- requires only two frames (low memory requirements)
- good value for the money:
  results are often superior to more complicated approaches

Disadvantages

- problems at locations where the local constancy assumption is violated:
  flow discontinuities and non-translatory motion (e.g. rotation)
- local method that does not allow to compute the flow field at all locations

The Spatiotemporal Approach of Lucas and Kanade

- In the Lucas–Kanade method one can replace the spatial local constancy assumption by a local spatiotemporal one.
- creates the same linear system

$$\begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * (f_x f_z) \\ -K_\rho * (f_y f_z) \end{pmatrix},$$

  but with a spatiotemporal Gaussian convolution that must be computed over the entire sequence.

- often more robust results than with spatial Lucas–Kanade

M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

The Spatiotemporal Approach of Bigün (1991)

- closely related to spatiotemporal Lucas–Kanade
- optic flow is regarded as orientation in the space–time domain and formulated as a principal component analysis problem of the structure tensor (cf. also Lecture 19).
- We search for the unit vector $\mathbf{w} = (w_1, w_2, w_3)^\top$ that minimises

$$E(\mathbf{w}) \;=\; \int\limits_{B_\rho(x_0,y_0,z_0)} (f_x w_1 + f_y w_2 + f_z w_3)^2 \, dx \, dy \, dz$$

- When renormalising the third component of the optimal $\mathbf{w}$ to 1, the first two components give the optic flow:

$$u = \frac{w_1}{w_3}, \qquad v = \frac{w_2}{w_3}.$$

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

- In contrast to (spatial or spatiotemporal) Lucas–Kanade, one does not search a 2-D flow vector with a least-square approach, but a 3-D vector, i.e. the time component is also optimised *(total least squares)*.
- Using the notation $\nabla_3 f := (f_x, f_y, f_z)^\top$ one wants to minimise

$$\begin{aligned}
E(\mathbf{w}) \;&:=\; \int\limits_{B_\rho} (\mathbf{w}^\top \nabla_3 f)^2 \, dx \, dy \, dz \\
&=\; \int\limits_{B_\rho} \mathbf{w}^\top \nabla_3 f \, \nabla_3 f^\top \mathbf{w} \, dx \, dy \, dz \\
&=\; \mathbf{w}^\top \Big(\int\limits_{B_\rho} \nabla_3 f \, \nabla_3 f^\top \, dx \, dy \, dz \Big) \, \mathbf{w}.
\end{aligned}$$

The Spatiotemporal Approach of Bigün (3)

- The desired vector $\mathbf{w}$ is the normalised eigenvector to the smallest eigenvalue of

$$\int_{B_\rho} \nabla_3 f \, \nabla_3 f^\top \, dx \, dy \, dz.$$

- Often the box filter is replaced by Gaussian convolution.
  Thus one performs a principal component analysis of the spatiotemporal structure tensor

$$\begin{aligned} J_\rho \quad &:= \quad K_\rho * (\nabla_3 f \, \nabla_3 f^\top) \\ &= \quad \begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) & K_\rho * (f_x f_z) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) & K_\rho * (f_y f_z) \\ K_\rho * (f_x f_z) & K_\rho * (f_y f_z) & K_\rho * (f_z^2) \end{pmatrix} \end{aligned}$$

The Spatiotemporal Approach of Bigün (4)

Flow Classification with the Eigenvalues of the Structure Tensor

- Let $\mu_1 \geq \mu_2 \geq \mu_3 \geq 0$ be the eigenvalues of J_ρ.

- $\mathrm{rank}(J_\rho) = 0$ (three vanishing eigenvalues):
  If $\mathrm{tr}\, J_\rho = j_{11} + j_{22} + j_{33} \leq \tau_1$, nothing can be said: All gradients are too close to 0.

- $\mathrm{rank}(J_\rho) = 3$ (no vanishing eigenvalues):
  If $\mu_3 \geq \tau_2$, then the assumption of a locally constant flow is violated. Either a flow discontinuity or noise dominate.

- $\mathrm{rank}(J_\rho) = 1$ (two vanishing eigenvalue):
  If $\mu_2 \leq \tau_3$, we have two low-contrast eigendirections. No unique flow exists (aperture problem). One can compute the normal flow

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \frac{-1}{f_x^2 + f_y^2} \begin{pmatrix} f_x f_z \\ f_y f_z \end{pmatrix}$$

- $\mathrm{rank}(J_\rho) = 2$ (one vanishing eigenvalue):
  In this case the optic flow results from the eigenvector $\mathbf{w}$ to the smallest eigenvalue μ_3. Normalising its third component to 1, the first two components give u and v.

Advantages

- high robustness with respect to noise
- good results for translatory motion
- eigenvalues of the spatiotemporal structure tensors provide detailed information on the optic flow

Disadvantages

- more complicated than Lucas–Kanade:
  numerical principal component analysis of a 3×3 matrix
  (suitable method: e.g. Jacobi transformations)
- problems at flow discontinuities and locations with non-translatory motion (e.g. rotation)
- local method that does not give full flow fields
- several threshold parameters

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

(a) Left: Image from the sphere sequence. **(b) Middle:** False colour representation of the optic flow using the Bigün method. **(c) Rigth:** Flow classification: black=no information (three small eigenvalues), dark grey=flow discontinuity or noise (three large eigenvalues), light grey=aperture problem (two small eigenvalues), white=full flow (one small eigenvalue). Author: J. Weickert (2001).

Summary

- Computing the optic flow is a key problem in computer vision.
- Assuming grey value constancy leads to the Optic Flow Constaint (OFC).
  It allows to compute the normal flow only (aperture problem).
  Computing the full flow requires additional assumptions.
- Lucas and Kanade assume a locally constant flow.
  This yields a linear system of equations with the spatial structure tensor as system matrix.
- The method of Bigün et al. estimates the flow as orientation in the spatiotemporal domain.
  It leads to a principal component analysis problem of the spatiotemporal stucture tensor.

Literature

- B. Lucas, T. Kanade: An iterative image registration technique with an application to stereo vision. In *Proc. Seventh International Joint Conference on Artificial Intelligence* (Vancouver, Canada, Aug. 1981), pp. 674–679.
  (original paper by Lucas and Kanade)
- J. Bigün, G.H. Granlund, J. Wiklund: Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, 775–790, 1991.
  (structure tensor method of Bigün et al.)
- B. Jähne: *Digital Image Processing*. Sixth Edition, Springer, Berlin, 2005.
  (Chapter 14 deals with optic flow.)
- J.L. Barron, D.J. Fleet, S.S. Beauchemin: Performance of optical flow techniques. *International Journal of Computer Vision*, Vol. 12, 43–77, 1994.
  (compares the performance of many optic flow methods)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Assignment C6 – Classroom Work

Problem 1 (Optic Flow Constraints)

In order to compute the optic flow from a sufficiently often continuously differentiable 2-D image sequence $f(x, y, t)$, many optic flow algorithms assume that the grey value of objects remains constant over time. However, in case of additive illumination changes it may make sense to formulate constancy assumptions on image features that are based on higher derivatives. Therefore let us assume for the moment that not the grey value, but the *magnitude of the gradient*, remains constant over time.

(a) Formulate the corresponding constancy assumption(s) without linearisation as well as the linearised Optic Flow constraint(s) that can be derived from it/them.

(b) Let us now assume that not the magnitude of the gradient but the *spatial gradient itself* remains constant over time. Formulate also for this case the corresponding constancy assumption(s) in its/their original form and as linearised constraint(s).

(c) Compare both cases regarding the presence of the aperture problem. Are there locations where the optic flow cannot be computed?

(*This problem shows how different constancy assumptions can by employed in the context of optic flow computation.*)

Assignment T6 – Theoretical Homework

Problem 1 (Mumford–Shah Cartoon Model) (6 points)

Let $\Omega_i,\ \Omega_j \subset \Omega$ denote two segments with mean u_i resp. u_j. Furthermore, let $\partial(\Omega_i,\ \Omega_j)$ denote the common boundary between Ω_i and Ω_j.
Show that for the Mumford–Shah cartoon model, merging these two regions results in the following change of energy:

$$E(K - \partial(\Omega_i,\ \Omega_j)) - E(K) \;=\; \frac{|\Omega_i| \cdot |\Omega_j|}{|\Omega_i| + |\Omega_j|}(u_i - u_j)^2 - \lambda\, l(\partial(\Omega_i,\ \Omega_j))\,.$$

(*This problem proves an essential property that is required for implementing the Mumford-Shah cartoon model.*)

Assignment T6 (2)

Problem 2 (The Method of Bigün *et al.*) (4 + 2 points)

The local differential optic flow method of Bigün *et al.* is based on thresholding the different eigenvalues of the spatiotemporal structure tensor $J_\rho(\nabla_3 f)$ and distinguishing four different cases.

(a) Which of the following matrices represent such a spatiotemporal structure tensor and which do not? Please specify in the first case, how many eigenvalues are non-zero, i.e. to which of the four cases of the method of Bigün *et al.* the corresponding tensor is related.

$$(i) \quad J_\rho = \begin{pmatrix} 5 & 2 & 3 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \qquad (ii) \quad J_\rho = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 6 & -3 \\ 1 & -3 & 7 \end{pmatrix}$$

$$(iii) \quad J_\rho = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -3 & -1 \\ 2 & -1 & 5 \end{pmatrix} \qquad (iv) \quad J_\rho = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(b) Why it is not desirable if all three eigenvalues of the spatiotemporal structure tensor $J_\rho(\nabla_3 f)$ are significantly larger than zero? Give an explanation.

(*This problem gives deeper insights into the idea of local differential optic flow methods.*)

Assignment T6 (3)

Problem 3 (Variational Optic Flow Methods) (3 + 2 + 3 points)

Let us consider a sufficiently often continuously differentiable 3-D image sequence $f(x, y, z, t)$. In order to compute the optic flow, let us furthermore assume that the grey value of objects in consecutive 3-D frames of this image sequence remain constant.

(a) Derive the linearised formulation for this 3-D grey value constancy assumption. What can you say about the aperture problem ?

(b) Embed the corresponding squared data term into a Horn-and-Schunck-like variational approach with suitable (quadratic) smoothness term. Write down the obtained energy functional.

(c) Compute the associated Euler–Lagrange equations. Are these equations linear or nonlinear?

(*This problem shows how to extend standard 2-D approaches to 3-D images sequences.*)

Deadline for submission: Tuesday, February 5, 10 am (before the lecture).