CS 565 Computer Vision

Nazar Khan

PUCIT

Lectures 15 and 16: Optic Flow

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 $(u(x,y,z),v(x,y,z))^T$
- Basically, where does pixel (x,y) move from frame z to frame z+1.
- Such correspondence problems are key problems in computer vision.

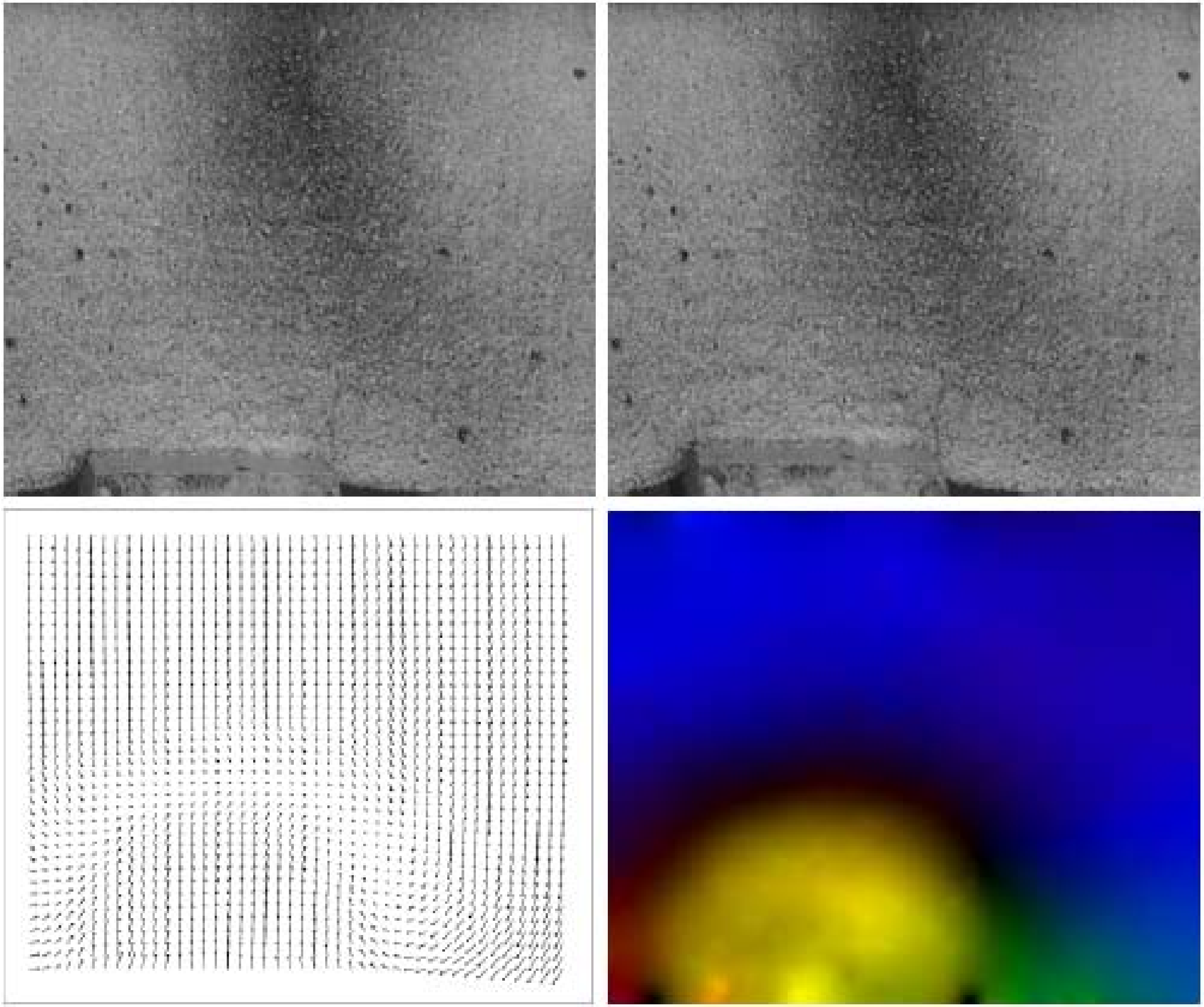
Similar Correspondence Problems

- 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

Introduction

What is Optic Flow Good for?

- recognition of moving pedestrians in driver assistant systems
- 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
- video processing, e.g. frame interpolation
- efficient video coding



Deformation analysis of plastic foam using an optic flow method. **Top left:** Frame 1 of a deformation sequence. **Top right:** Frame 2. **Bottom left:** Vector plot of the displacement field. **Bottom right:** Colour-coded displacement field. Author: J. Weickert (2002).

Grey Value Constancy Assumption

- Corresponding image structures should have the same grey value.
- Thus, the optic flow between frame z and z + 1 satisfies $f(x+u, y+v, z+1) = f(x, y, z)$.
- Unfortunately the unknown flow field $(u, v)^T$ is not directly accessible.
  - This problem is similar to the Harris corner detection formulation where direction d was also not accessible. **(How did we get around that problem?)**

Linearisation by Taylor Expansion

- Let us assume that (u, v) is small and f varies slowly.
- Then a Taylor expansion around (x, y, z) gives a good approximation

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

(H.W. Prove this.)

where subscripts denote partial derivatives.

- This yields the linearised optic flow constraint (OFC)

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

where the unknown flow field $(u, v)^T$ is directly accessible.

Assumptions

We have made 2 assumptions so far:

1. Grey value constancy
2. Linearised OFC

How Realistic are These Assumptions?

- The grey value constancy assumption is often surprisingly realistic:
  - Many illumination changes happen very slowly, i.e. over many frames.
  - More complicated models exist that take into account illumination changes.

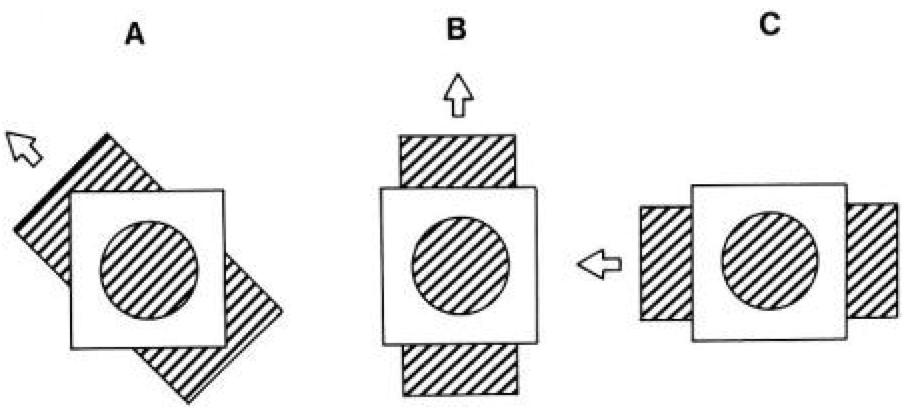
- The linearisation assumption is violated more frequently:
  - Conventional video cameras often suffer from temporal undersampling (produce displacements over several pixels) while Taylor expansion is accurate only for small displacements.
  - Remedies:
    - use original OFC without linearisation (model becomes more difficult)
    - spatial downsampling (after lowpass filtering!) **(H.W. How will this help?)**

The Aperture Problem

- The OFC $f_x u + f_y v + f_z = 0$ is one equation in two unknowns u, v. Thus, it cannot have a unique solution.
- The OFC specifies only the flow component parallel to the spatial gradient $\nabla f = (f_x, f_y)^T$:

$$0 = f_x u + f_y v + f_z = [u \;\; v]\nabla f + f_z$$

- This sheds more light on the non-uniqueness problem:
  - Adding arbitrary flow components orthogonal to ∇f does not violate the OFC. This is called aperture problem.



Within the viewing circle (aperture), movements in A, B and C will **appear** the same even though the **real** movements are all different.

http://stoomey.wordpress.com/2008/04/18/20/

The Aperture Problem

- Additional assumptions are necessary to get a unique solution.
- Specifying different additional constraints leads to different methods.
- Let us first analyse the flow component along ∇f.

The Normal Flow

- Expressing the flow vector $(u, v)^T$ in terms of the basis vectors $n=\nabla f/|\nabla f|$ and $t=\nabla f^{\perp}/|\nabla f|$ gives the flow normal and tangential to the edge of f:

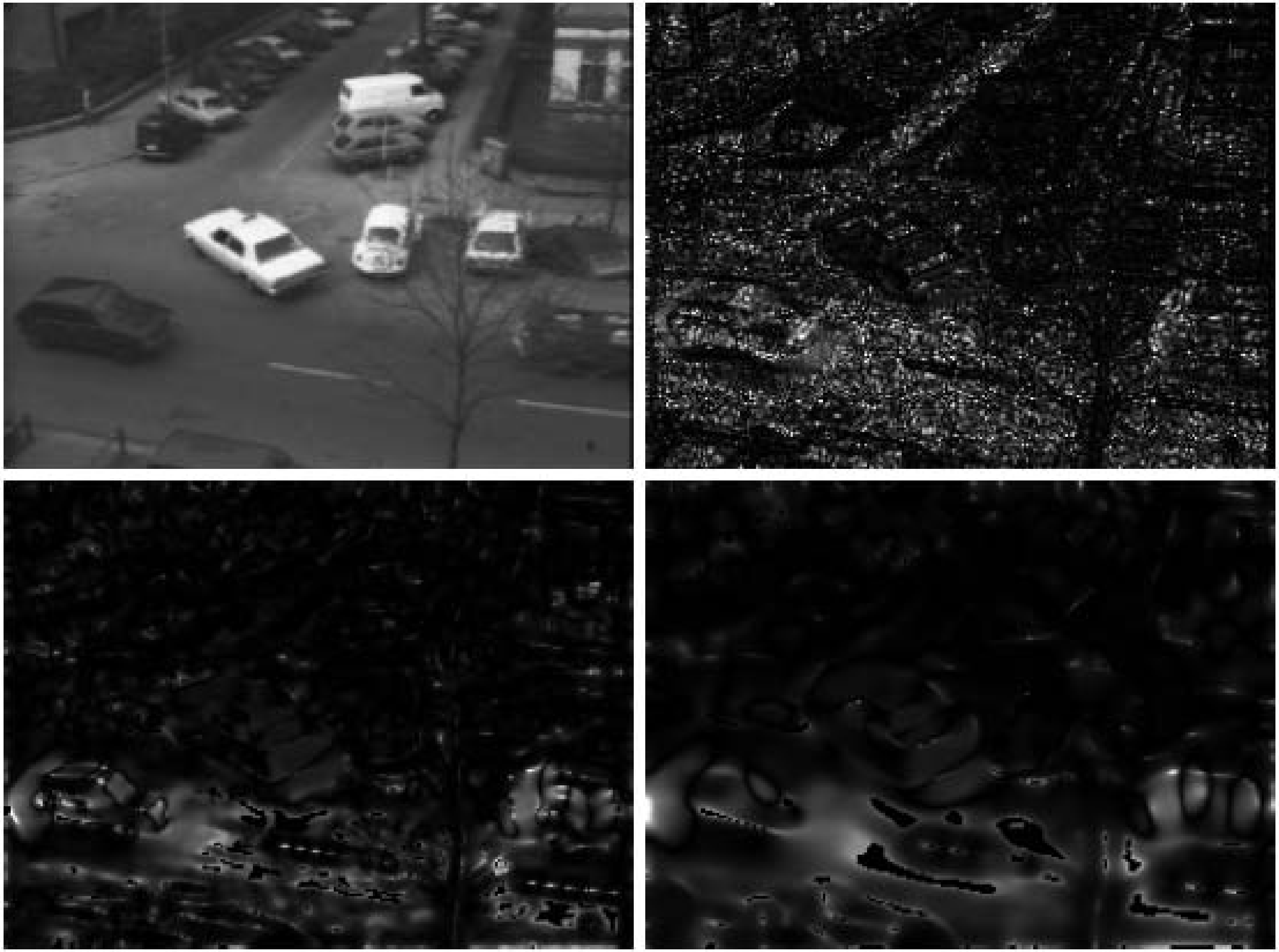
$$(u,v)^T = (u, v)\nabla f/|\nabla f| \; \nabla f/|\nabla f| + (u, v)\; \nabla f^{\perp}/|\nabla f| \; \nabla f^{\perp}/|\nabla f|$$
$$=: (u_n,v_n)^T + (u_t,v_t)^T.$$

- The OFC yields $(u, v)\nabla f = -f_z$, and the normal flow becomes

$$(u_n,v_n)^T = - f_z/|\nabla f|. \; \nabla f/|\nabla f| = -1/(f_x^2+f_y^2)\, (f_xf_z,f_yf_z)^T$$

- The normal flow is the only flow that can be computed from the OFC without additional constraints.
  - Unfortunately, it gives poor results.

Hamburg Taxi Sequence



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

The Spatial Approach of Lucas and Kanade

- 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 disk-shaped neighbourhood $B(x_0, y_0)$ of radius ρ.
- least squares model: flow in (x_0, y_0) minimises the local energy

$$E(u,v) = \frac{1}{2} \int_{B_\rho(x0,y0)} (f_x u + f_y v + f_z)^2 \, dx \, dy$$

The Spatial Approach of Lucas and Kanade

- least squares model: flow in (x_0, y_0) minimises the local energy

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

- Computing partial derivatives and equating to 0

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

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

The Spatial Approach of Lucas and Kanade

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

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

- Multiplying both sides by $1/|B_\rho|$ does not change the linear system.
- $1/|B_\rho|$ can be multiplied with each integral on both sides.
- So the integrals convert to averages.
- Averaging can be replaced by weighted averaging.
- One form of weighted averaging is Gaussian smoothing.

The Spatial Approach of Lucas and Kanade

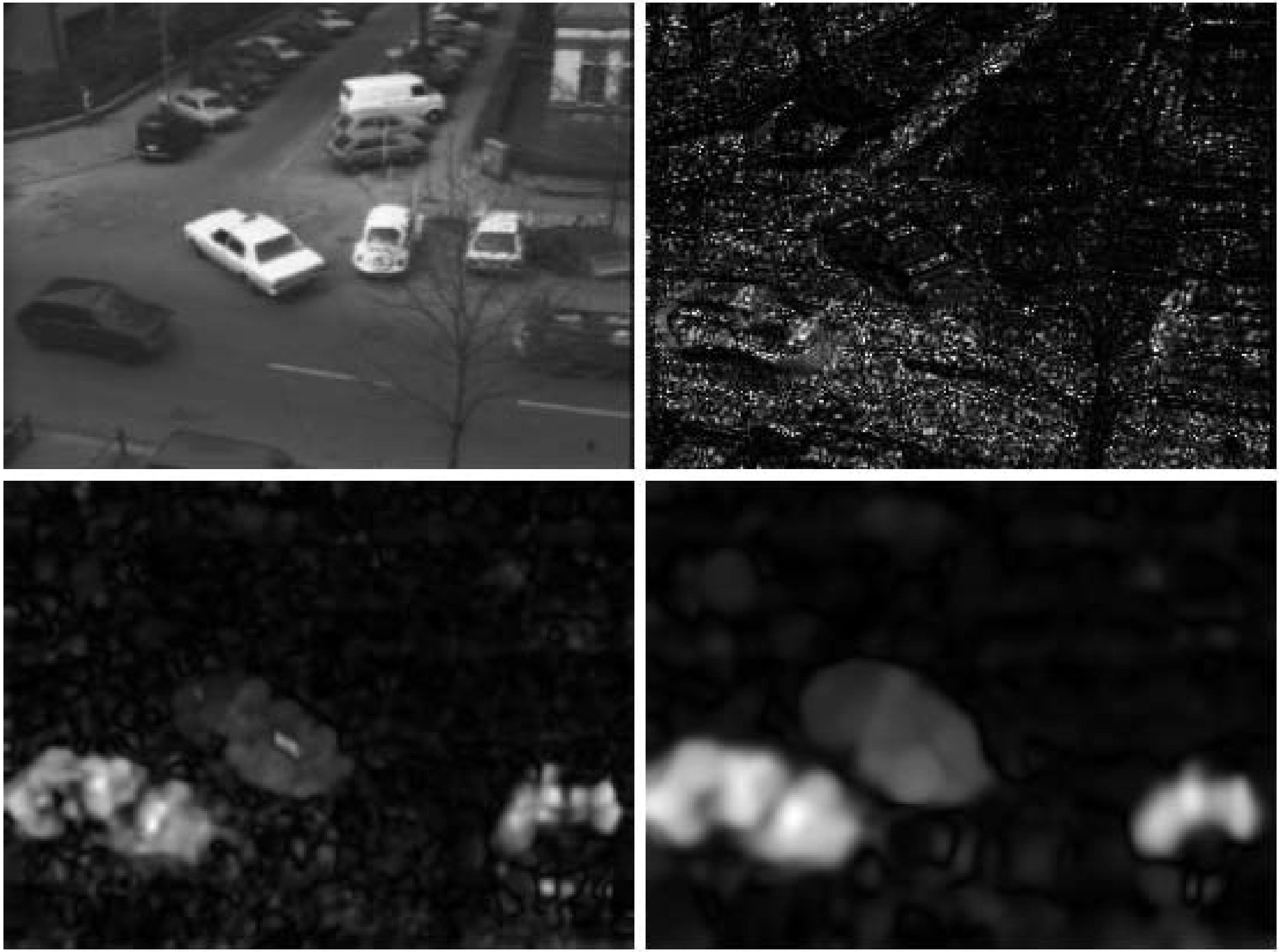
- Often one replaces the box filter with a “hard” window B(x, y) by a “smooth” convolution with a Gaussian K_ρ:

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

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

The Spatial Approach of Lucas and Kanade

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



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

The Spatial Approach of Lucas and Kanade

When Does the Linear System Have No Unique Solution?

- rank(J) = 0 (two vanishing eigenvalues):

  Happens if the spatial gradient vanishes in the entire neighbourhood.

  Nothing can be said in this case.

  Simple criterion: trace (J) = $j_{1,1} + j_{2,2} \leq \varepsilon$.

  (Remember that J is positive semidefinite)

The Spatial Approach of Lucas and Kanade

When Does the Linear System Have No Unique Solution?

- rank(J) = 1 (one vanishing eigenvalue):

  Happens if we have the same (nonvanishing) spatial gradient within the entire neighbourhood.

  Then both equations are linearly dependent (infinitely many solutions).

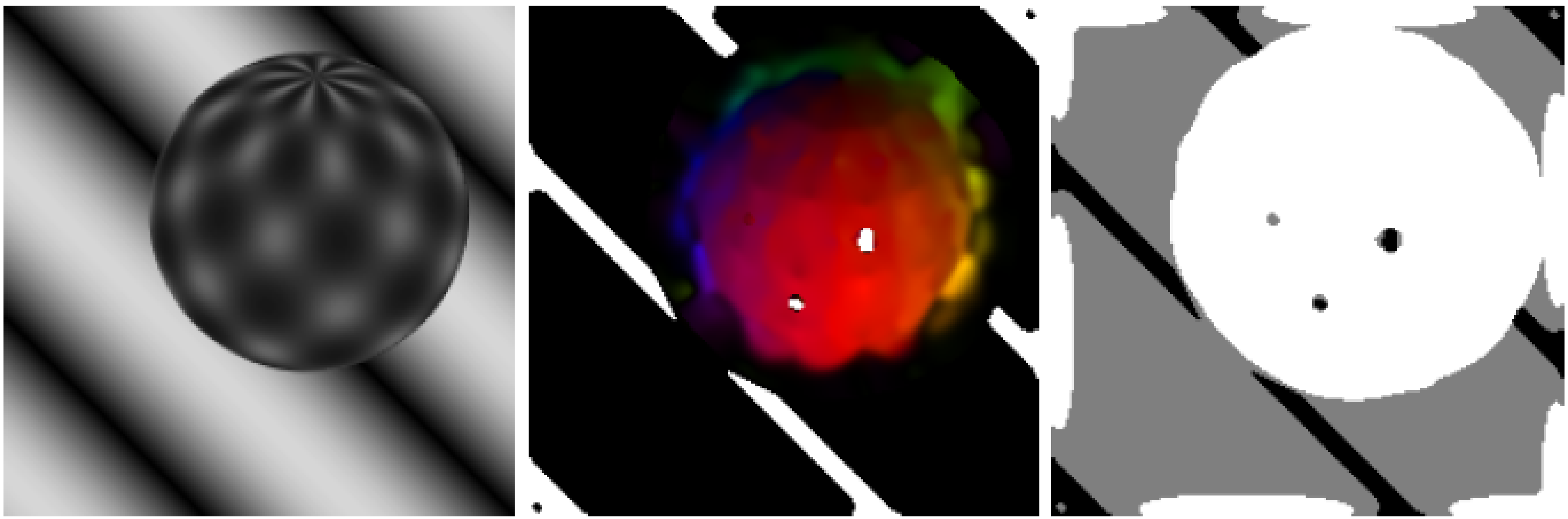
  Simple criterion: det (J) = $j_{1,1}\, j_{2,2} - j_{1,2}^2 \leq \varepsilon$ (while trace(J) > ε).

  In this case the aperture problem persists.

  One can only compute the normal flow

$$(u_n,v_n)^T = -1/(f_x^2+f_y^2)\ (f_xf_z,f_yf_z)^T$$

The Spatial Approach of Lucas and Kanade



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

The Spatial Approach of Lucas and Kanade

Advantages

- simple and fast method
- requires only two frames (low memory requirements)
- good value for money: results 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 Biguen et al.

- Optic flow is regarded as orientation in the space–time domain and formulated as a principal component analysis problem of the structure tensor.
- We search for the direction with the least grey value changes within a 3-D ball-shaped neighbourhood $B(x_0,y_0,z_0)$ of radius ρ.

The Spatiotemporal Approach of Biguen et al.

- It is given by the unit vector $w=(w_1,w_2,w_3)^T$ that minimises

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

- When re-normalising the third component of the optimal w to 1, the first two components give the optic flow:

$$u = w_1/w_3, \qquad v = w_2/w_3$$

The Spatiotemporal Approach of Biguen et al.

- Using the spatiotemporal gradient notation $\nabla_3 f := (f_x, f_y, f_z)^\top$ one minimises

$$E(w) := \int_{B_\rho} (w^\top \nabla_3 f)^2 \, dx \, dy \, dz$$

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

$$= w^\top \Big(\int_{B_\rho} \nabla_3 f \, \nabla_3 f^\top \, dx \, dy \, dz \Big) \, w$$

with the constraint $||w|| = 1$

The Spatiotemporal Approach of Biguen et al.

- The desired vector w is the normalised eigenvector to the smallest eigenvalue of

$$\int_{B_\rho} \boldsymbol{\nabla}_3 f \, \boldsymbol{\nabla}_3 f^\top \, dx \, dy \, dz$$

- Summation in region B_ρ can be replaced by Gaussian convolution. Leads to a principal component analysis of the spatiotemporal structure tensor $J_\rho := K_\rho * (\boldsymbol{\nabla}_3 f \, \boldsymbol{\nabla}_3 f^\top)$

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

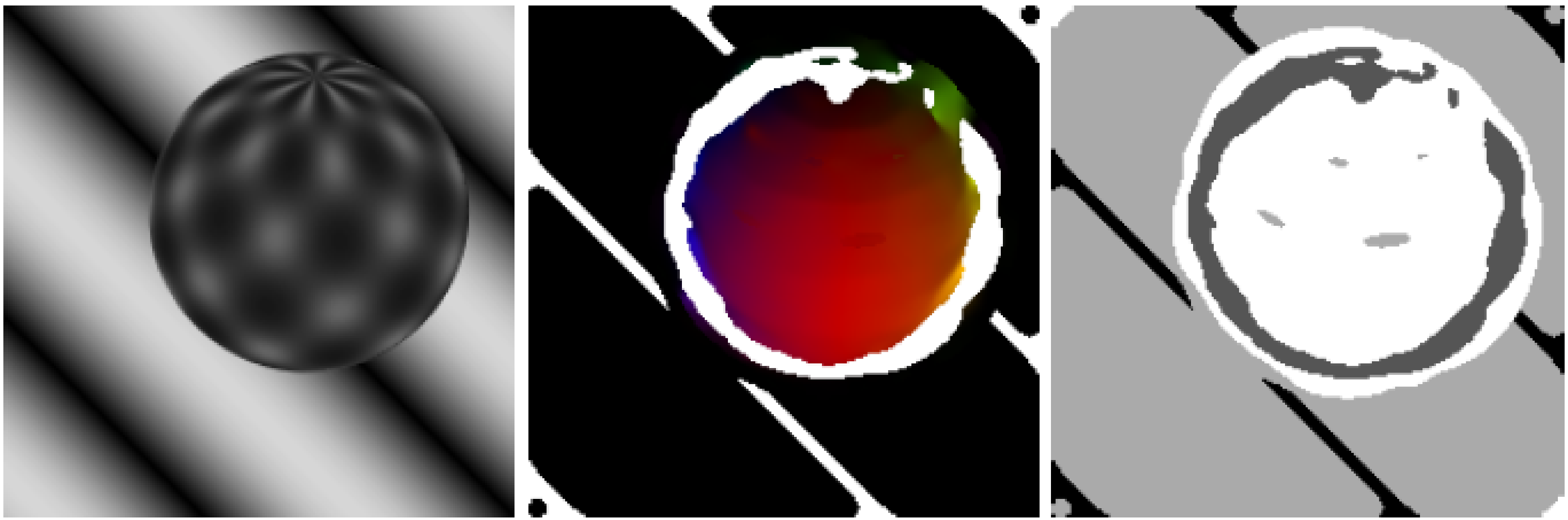
The Spatiotemporal Approach of Biguen et al.

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.

- rank(J) = 0 (three vanishing eigenvalues):
  If tr J = $j_{1,1} + j_{2,2} + j_{3,3} \leq \tau_1$, nothing can be said: The gradients are too small.
- rank(J) = 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 dominates.
- rank(J) = 1 (two vanishing eigenvalues):
  If $\mu_2 \leq \tau_3$, we have two low-contrast eigendirections. No unique flow exists (aperture problem). One can compute the normal flow only.
- rank(J) = 2 (one vanishing eigenvalue):
  In this case the optic flow results from the eigenvector w to the smallest eigenvalue μ_3. Normalising its third component to 1, the first two components give u and v.

The Spatiotemporal Approach of Biguen et al.



Left: Image from the sphere sequence. **Middle:** False colour representation of the optic flow using the Bigün method. **Right:** 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).

The Spatiotemporal Approach of Biguen et al.

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
- 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

Summary of Local Optic Flow Methods

- Assuming grey value constancy leads to the Optic Flow Constraint (OFC).
  - It allows to compute the normal flow only (aperture problem).
  - Computing the full flow requires additional assumptions.
- Lucas and Kanade assume a <u>locally constant flow</u> (in 2D).
  - This yields a linear system of equations with the spatial structure tensor as system matrix.

Summary of Local Optic Flow Methods

- The method of Biguen et al. estimates the flow as orientation in the spatiotemporal domain.
  - It leads to a principal component analysis problem of the spatiotemporal structure tensor.
- Both are local methods that do not compute the flow at every pixel. That is, the flow field is not dense.

Variational Method of Horn and Schunck

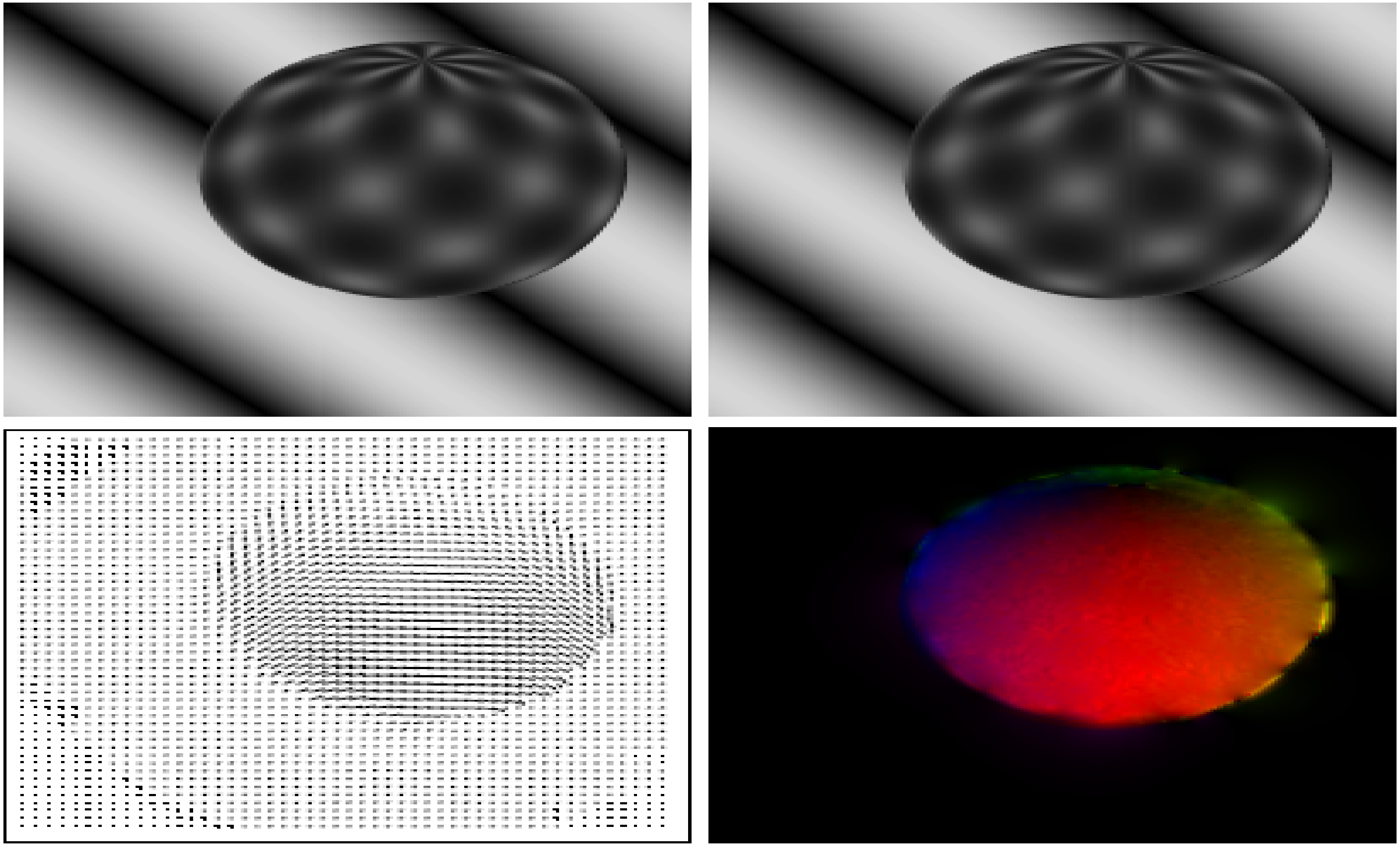
- At some given time z the optic flow field is determined as minimising the function $(u(x, y), v(x, y))^T$ of the energy functional

$$E(u,v) := \frac{1}{2}\int_\Omega \Big(\underbrace{\big(f_x u + f_y v + f_z\big)^2}_{\text{data term}} + \alpha \underbrace{\big(|\nabla u|^2 + |\nabla v|^2\big)}_{\text{smoothness term}} \Big)\, dx\, dy$$

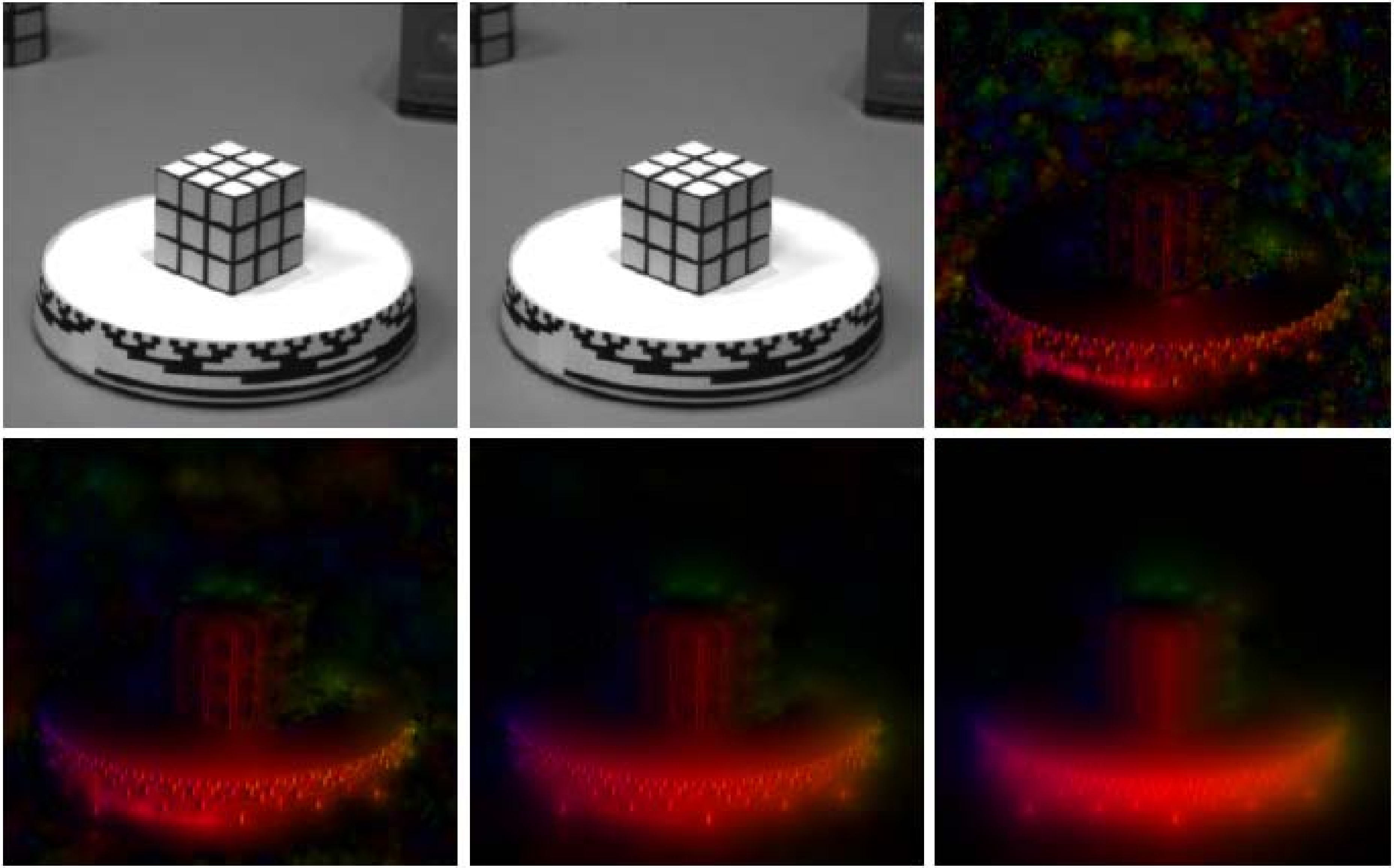
- Has a unique solution that depends continuously on the image data.

Variational Method of Horn and Schunck

- Regularisation parameter $\alpha > 0$ determines smoothness of the flow field:
  - $\alpha \rightarrow 0$ yields the normal flow.
  - The larger α, the smoother the flow field.



Optic flow computation using the Horn–Schunck method. **Top left**: Frame 10 of a synthetic image sequence. **Top right:** Frame 11. **Bottom left:** Optic flow, vector plot. **Bottom right**: Optic flow, colour-coded. Author: J. Weickert (2000).



Influence of the regularisation parameter. **Top left:** Frame 10 of the rotating cube sequence. **Top middle:** Frame 11. **Top right:** Optic flow, $\alpha = 1$. **Bottom left:** $\alpha = 10$. **Bottom middle:** $\alpha = 100$. **Bottom right:** $\alpha = 1000$. Author: J. Weickert (2000).

Variational Method of Horn and Schunck

Main advantage

- Dense flow fields due to filling-in effect:
  - At locations, where no reliable flow estimation is possible (small $|\nabla f|$), the smoothness term dominates over the data term.
- This propagates data from the neighbourhood.
- No additional threshold parameters necessary

How to solve for the flow field (u,v)?

Step 1: Going to the Euler-Lagrange Equations

Important Result from Calculus of Variations

Minimiser of the energy functional

$$E(u,v) := \int_\Omega F(x,y,u,v,u_x,u_y,v_x,v_y)\,dx\,dy$$

satisfies the Euler–Lagrange equations

$$\begin{aligned}\partial_x F_{u_x} + \partial_y F_{u_y} - F_u &= 0,\\ \partial_x F_{v_x} + \partial_y F_{v_y} - F_v &= 0\end{aligned}$$

with boundary conditions

$$n^\top \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} = 0, \qquad n^\top \begin{pmatrix} F_{v_x} \\ F_{v_y} \end{pmatrix} = 0$$

How to solve for the flow field (u,v)?

Application to Our Problem

The integrand

$$F = \frac{1}{2}\left(f_x u + f_y v + f_z\right)^2 + \frac{\alpha}{2}\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right)$$

has the partial derivatives

$$\begin{aligned}
F_u &= f_x(f_x u + f_y v + f_z), \\
F_v &= f_y(f_x u + f_y v + f_z), \\
F_{u_x} &= \alpha\, u_x, \\
F_{u_y} &= \alpha\, u_y, \\
F_{v_x} &= \alpha\, v_x, \\
F_{v_y} &= \alpha\, v_y.
\end{aligned}$$

How to solve for the flow field (u,v)?

This yields the Euler–Lagrange equations

$$\begin{aligned} \alpha\,\Delta u \;-\; f_x\,(f_x u + f_y v + f_z) &= 0, \\ \alpha\,\Delta v \;-\; f_y\,(f_x u + f_y v + f_z) &= 0. \end{aligned}$$

After division by α, the boundary conditions are given by

$$\begin{aligned} 0 &= \boldsymbol{n}^\top \boldsymbol{\nabla} u = \partial_{\boldsymbol{n}} u, \\ 0 &= \boldsymbol{n}^\top \boldsymbol{\nabla} v = \partial_{\boldsymbol{n}} v. \end{aligned}$$

How to solve for the flow field (u,v)?

Step 2: Discretisation

- Approximate required first and second order derivatives using simple difference operators.
- Yields the difference equations

$$0 = \frac{\alpha}{h^2} \sum_{j \in \mathcal{N}(i)} (u_j - u_i) - f_{xi} \, (f_{xi} \, u_i + f_{yi} \, v_i + f_{zi}),$$

$$0 = \frac{\alpha}{h^2} \sum_{j \in \mathcal{N}(i)} (v_j - v_i) - f_{yi} \, (f_{xi} \, u_i + f_{yi} \, v_i + f_{zi})$$

  for all pixels (i=1,…,N) where h is the grid size (usually 1).

- Can be written as a sparse but <u>very large</u> linear system Bx=d.
  - Size of B will be 69GB for a 256x256 image!

How to solve for the flow field (u,v)?

Step 3: Solving the Linear System

- Jacobi Method: Iterative way of solving Bx=d
  1. Let B=D−N with a diagonal matrix D and a remainder N.
  2. Then the problem Dx = Nx + d is solved iteratively using

$$x^{(k+1)} = D^{-1}(Nx^{(k)} + d)$$

- low computational effort per iteration if B is sparse:
  – 1 matrix–vector product, 1 vector addition, 1 vector scaling
- only small additional memory requirement: vector $x^{(k)}$
- well-suited for parallel computing
- residue $r^{(k)} := Bx^{(k)} - d$ allows simple stopping criterion:

$$\text{stop if } |r^{(k)}| / |r^{(0)}| < \varepsilon$$

How to solve for the flow field (u,v)?

- All of the above boils down to a very simple iterative scheme

$$u_i^{(k+1)} = \frac{\frac{\alpha}{h^2} \sum_{j \in \mathcal{N}(i)} u_j^{(k)} - f_{xi} \left(f_{yi} \, v_i^{(k)} + f_{zi} \right)}{\frac{\alpha}{h^2} |\mathcal{N}(i)| + f_{xi}^2},$$

$$v_i^{(k+1)} = \frac{\frac{\alpha}{h^2} \sum_{j \in \mathcal{N}(i)} v_j^{(k)} - f_{yi} \left(f_{xi} \, u_i^{(k)} + f_{zi} \right)}{\frac{\alpha}{h^2} |\mathcal{N}(i)| + f_{yi}^2}$$

  with k = 0, 1, 2, ... and an arbitrary initialisation (e.g. null vector).
- All of you can implement this easily! **(Assignment 5)**

Flow estimate at pixel *j* at iteration *k*

Flow estimate at pixel *i* at iteration *k*

Flow estimate at pixel *i* at iteration *k+1*

$$u_i^{(k+1)} = \frac{\frac{\alpha}{h^2} \sum_{j \in \mathcal{N}(i)} u_j^{(k)} - f_{xi} \left(f_{yi} \, v_i^{(k)} + f_{zi} \right)}{\frac{\alpha}{h^2} |\mathcal{N}(i)| + f_{xi}^2},$$

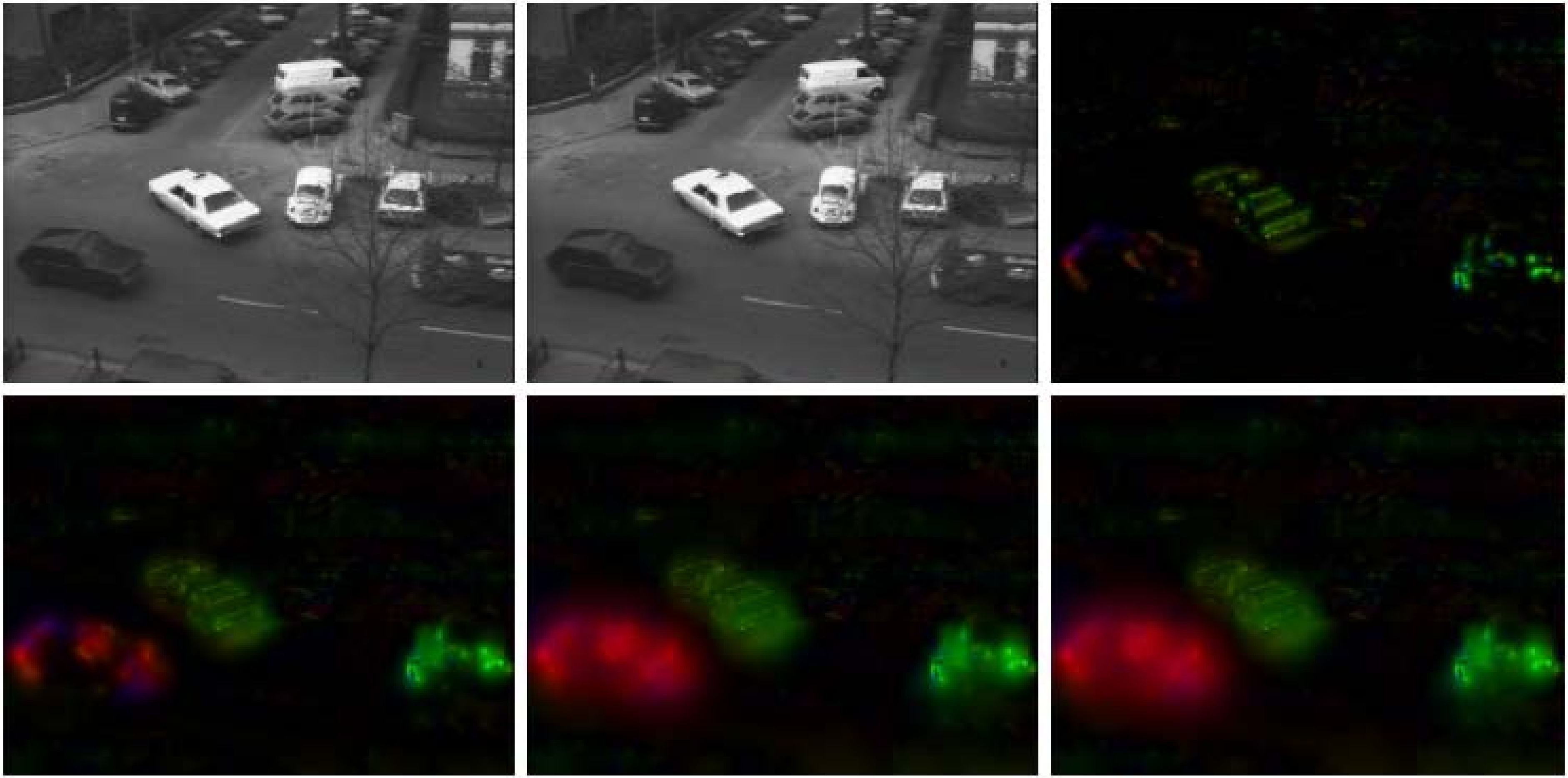
$$v_i^{(k+1)} = \frac{\frac{\alpha}{h^2} \sum_{j \in \mathcal{N}(i)} v_j^{(k)} - f_{yi} \left(f_{xi} \, u_i^{(k)} + f_{zi} \right)}{\frac{\alpha}{h^2} |\mathcal{N}(i)| + f_{yi}^2}$$

h=grid distance (usually $h=1$)

α=smoothness parameter

$N(i)$=set of neighboring pixels of pixel i

f_{xi}, f_{yi}, f_{zi} = spatial and temporal gradients at pixel i.



Influence of the number of Jacobi iterations. **Top left:** Frame 10 of the Hamburg taxi sequence. **Top middle:** Frame 11. **Top right:** Optic flow after 10 iterations. **Bottom left:** 100 iterations. **Bottom middle:** 1000 iterations. **Bottom right:** 10000 iterations. Author: J. Weickert (2000).

Summary of Global Optic Flow Methods

- Variational methods for computing the optic flow are global methods.
- Create dense flow fields by filling-in
- Model assumptions of the variational Horn and Schunck approach:
    1. grey value constancy,
    2. smoothness of the flow field
- Mathematically well-founded

Summary of Global Optic Flow Methods

- Minimising the energy functional leads to coupled differential equations.
- Discretisation creates a large, sparse linear system of equations.
  – can be solved iteratively, e.g. using the Jacobi method
- Variational methods can be extended and generalised in numerous ways, both with respect to models and to algorithms.