Lecture 3:

Optic Flow I
Local Differential Methods, Parameterisation Models

Contents

Optic Flow Estimation

What is the Optic Flow Problem?

Given: Two consecutive images of an image sequence (→ one camera, different time).
Wanted: Motion field between both frames (arbitrary motion possible).

Example for Optic Flow Estimation

- Human Motion Analysis
  (*Bruhn - Unimagazin CAMPUS 2007*)

Dancing Lecturer Sequence. **(a) Left:** Frame 1. **(b) Middle:** Optic Flow. **(c) Right:** Frame 2.

Fields of Application

- Robot navigation
- Driver assistance systems
- Video compression
- Surveillance/Tracking
- Superresolution/Deinterlacing

Challenges in Video Sequences

- Arbitrary motion
- Varying illumination (e.g. automatic brightness adjustment)
- Shadows and shading
- Camera noise
- Occlusions
- Appearing and disappearing objects

Quality Measures

How to Measure the Quality of Optic Flow Results if a Ground Truth is Given?

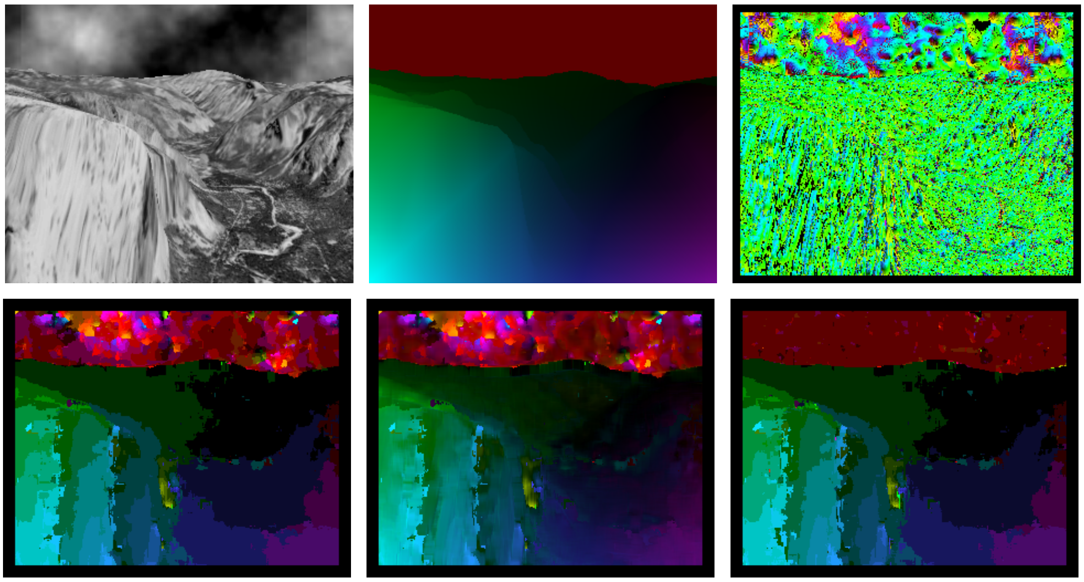
- *Idea:* Quantify difference between estimated flow field $\mathbf{u}^{\mathrm{e}}$ and ground truth $\mathbf{u}^{\mathrm{t}}$
- Spatiotemporal Average Angular Error (AAE)

$$AAE = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \arccos \left(\frac{u^{\mathrm{t}}_{i,j} u^{\mathrm{e}}_{i,j} + v^{\mathrm{t}}_{i,j} v^{\mathrm{e}}_{i,j} + 1}{\sqrt{{u^{\mathrm{t}}_{i,j}}^2 + {v^{\mathrm{t}}_{i,j}}^2 + 1} \sqrt{{u^{\mathrm{e}}_{i,j}}^2 + {v^{\mathrm{e}}_{i,j}}^2 + 1}} \right)$$

- Average Absolute Difference Error (AADE)

$$AADE = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \sqrt{\left(u^{\mathrm{t}}_{i,j} - u^{\mathrm{e}}_{i,j}\right)^2 + \left(v^{\mathrm{t}}_{i,j} - v^{\mathrm{e}}_{i,j}\right)^2}$$

How Good Are Block Matching Methods for Computing the Optic Flow?



Results for the Yosemite Sequence with clouds (L. Quam). **(a) Upper Left:** Frame 8. **(b) Upper Center:** Ground truth. **(c) Upper Right:** Basic approach ($d = 7$). **(d) Lower Left:** SSD ($m = 4$, $d = 7$). **(e) Lower Center:** SSD + Subpixel ($m = 4$, $d = 7$) **(f) Lower Right:** NCC ($m = 4$, $d = 7$).

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 27 28 29 30 31 32 33 34 35 36 37 38

Comparison in Terms of the Average Angular Error (AAE)

- Qualitative Evaluation for the Yosemite Sequence with Clouds

Technique	AAE
Basic Approach	72.89°
Zero Flow Field	55.22°
Block Matching (SSD)	24.44°
Block Matching (SAD)	24.40°
Normalised Cross Correlation (NCC)	21.84°
Block Matching + Subpixel (SSD)	21.46°

- *Conclusion:* Block matching methods offer a very poor performance if used for computing the optic flow. We have to consider some other approaches.
- In this lecture we present methods that allow for an AAE of less than 10.00°

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 27 28 29 30 31 32 33 34 35 36 37 38

Continuous Modelling

From Discrete to Continuous Modelling

- So far we have assumed that our images $\mathbf{f}$ and $\mathbf{g}$ are discrete
- *Idea:* Let us now consider these images as instances of a continuous image sequence $f_0(x, y, t)$, where $(x, y)^\top \in \Omega$ is the location and $t \in \mathbb{R}_0^+$ denotes time

$f_0(x, y, t)$

$f_0(x, y, t+1)$

- For fixed t two consecutive frames are given by $f_0(x, y, t)$ and $f_0(x, y, t+1)$

Gaussian Presmoothing Step

- *Idea*: Convolve the initial image sequence f_0 with a Gaussian K_σ of mean $\mu = 0$ and standard deviation σ

$$f(x, y, t) = K_\sigma * f_0(x, y, t)$$

  - reduces the influence of noise and outliers while preserving the mean value
  - image sequence becomes infinitely many times differentiable, i.e. $f \in \mathcal{C}^\infty$

$\xrightarrow{K_\sigma *}$

- Both aspects are important for differential methods that rely on the computation of image derivatives. Such methods are the main focus of this class.

Continuous Modelling (3)

The Grey Value Constancy Assumption

- If we denote the searched displacements in x- and y-direction by $u(x, y, t)$ and $v(x, y, t)$, respectively, the *continuous grey value constancy assumption* reads

$$f(x, y, t) - f(x + u, y + v, t + 1) = 0\,.$$

The Linearised Grey Value Constancy Assumption

- If u and v are small and f is sufficiently smooth, one may *linearise* this constancy assumption via a first-order Taylor expansion around the point $(x, y, t)^\top$:

$$f_x u + f_y v + f_t = 0\,.$$

  This constraint is the **brightness constancy constraint equation (BCCE)**. In general such constraints on the flow field are called optic flow constraints (OFCs).

Continuous Modelling (4)

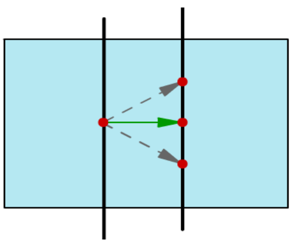
M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38

The Aperture Problem

- The BCCE provides only one equation for determining two unknowns
- Ill-posed problem with infinitely many solutions
- Only the flow component in direction of the image gradient can be computed, the so-called **normal flow**:

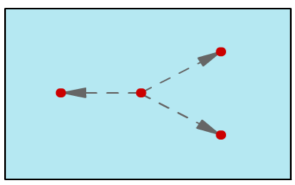
$$(u, v)_\mathrm{n}^\top = \frac{-f_t}{|\nabla f|}\frac{\nabla f}{|\nabla f|}\,.$$

- This problem is referred to as the **aperture problem**. It can be illustrated as



Case I

$|\nabla f| \neq 0 \rightarrow$ Aperture problem

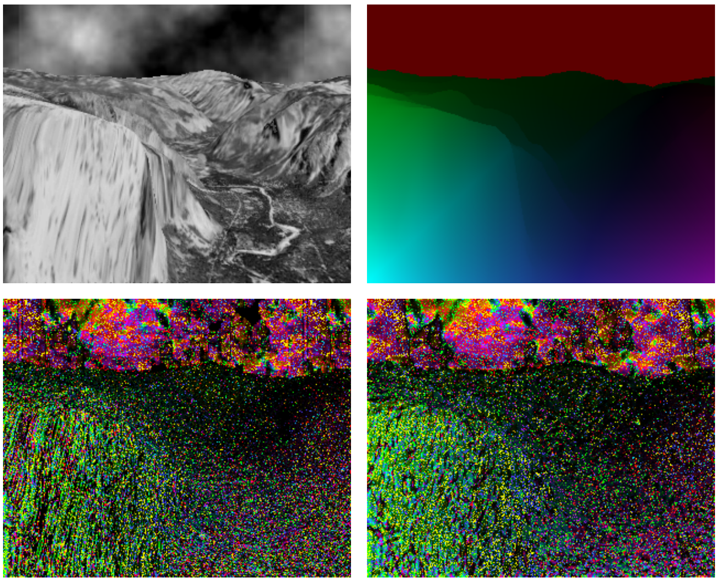


Case II

$|\nabla f| = 0 \rightarrow$ No estimation possible

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 27 28 29 30 31 32 33 34 35 36 37 38

How Accurate is the Normal Flow?



Results for the Yosemite Sequence with clouds (L. Quam). **(a) Upper Left:** Frame 8. **(b) Upper Right:** Ground truth. **(c) Lower Left:** Normal flow ($\sigma = 0$) with AAE=55.56° . **(d) Lower Right:** Normal flow ($\sigma = 1.4$) with AAE=50.56°.

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 27 28 29 30 31 32 33 34 35 36 37 38

The Spatial Method of Lucas and Kanade

How Can One Overcome the Aperture Problem?

- *Idea:* Assume flow to be locally constant and use neighbourhood information
- The method of Lucas and Kanade minimises the local energy
  (Lucas/Kanade 1982, Baker/Matthews 2004)

$$E((u(x_0,y_0), v(x_0,y_0)) = \int_{\mathcal{N}_\rho(x_0,y_0)} \Big(f_x(x,y,t)\,u + f_y(x,y,t)\,v + f_t(x,y,t)\Big)^2 \,dx\,dy\,.$$

  - $\mathcal{N}_\rho(x_0,y_0)$: spatial neighbourhood window of size ρ around $(x_0,y_0)^\top$
  - can be understood as a **least squares fit** w.r.t. to u and v
  - similar to block matching approaches with quadratic cost function
- Advantage over discrete methods: The solution can be computed **explicitly**

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 27 28 29 30 31 32 33 34 35 36 37 38

Minimisation

- The minimum requires the derivatives w.r.t. u and v to be zero, i.e.

$$0 \stackrel{!}{=} \frac{\partial E}{\partial u} = 2 \int_{\mathcal{N}_\rho(x_0,y_0)} f_x\,(f_x u + f_y v + f_t)\;dx\,dy\;,$$

$$0 \stackrel{!}{=} \frac{\partial E}{\partial v} = 2 \int_{\mathcal{N}_\rho(x_0,y_0)} f_y\,(f_x u + f_y v + f_t)\;dx\,dy\;.$$

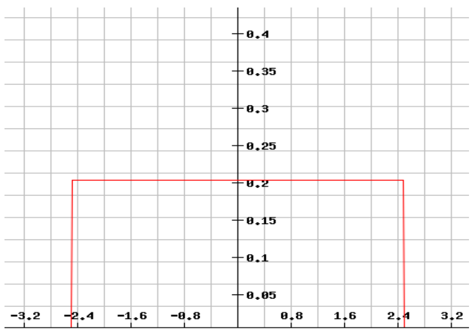
- Since u and v are assumed to be constant, they can be moved out of the integral. Then one obtains the following 2×2 linear systems of equations:

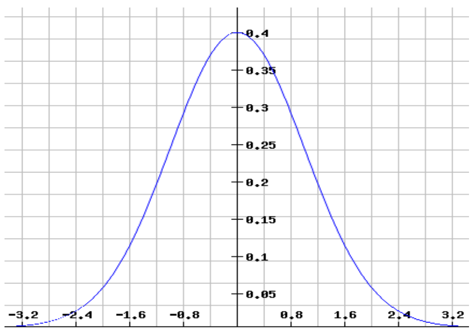
$$\begin{pmatrix} \int_{\mathcal{N}_\rho} f_x f_x\;dx\,dy & \int_{\mathcal{N}_\rho} f_x f_y\;dx\,dy \\ \int_{\mathcal{N}_\rho} f_x f_y\;dx\,dy & \int_{\mathcal{N}_\rho} f_y f_y\;dx\,dy \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\int_{\mathcal{N}_\rho} f_x f_t\;dx\,dy \\ -\int_{\mathcal{N}_\rho} f_y f_t\;dx\,dy \end{pmatrix}\;.$$

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 27 28 29 30 31 32 33 34 35 36 37 38

Smooth Windowing

- Replace hard window $\mathcal{N}_\rho$ by convolution with smooth Gaussian K_ρ





- The linear 2×2 system of equations then becomes

$$\underbrace{\begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} -K_\rho * (f_x f_t) \\ -K_\rho * (f_y f_t) \end{pmatrix}}_{\mathbf{b}}\;,$$

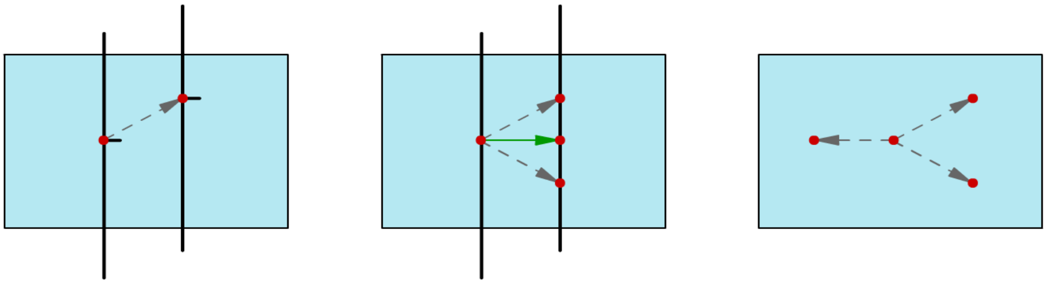
where the matrix A can be identified as the structure tensor (c.f. Lecture 2).

The Method of Lucas and Kanade (4)

Local Solvability and Uniqueness of the Solution

- The rank of the structure tensor decides on the local solvability of $A\mathbf{x} = \mathbf{b}$:
  - Rank 2 ($\lambda_1 \gg 0,\ \lambda_2 \gg 0$) – variation in both directions $\rightarrow$ full flow
  - Rank 1 ($\lambda_1 \gg 0,\ \lambda_2 \approx 0$) – variation in one direction $\rightarrow$ normal flow
  - Rank 0 ($\lambda_1 \approx 0,\ \lambda_2 \approx 0$) – no local structure $\rightarrow$ no flow

  Only if A has rank 2 a unique solution exists. Otherwise the aperture problem persists and infinitely many solutions are valid.



Case I — Rank 2 $\longrightarrow$ Full flow

Case II — Rank 1 $\longrightarrow$ Normal flow

Case III — Rank 0 $\longrightarrow$ No flow

MIA 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 27 28 29 30 31 32 33 34 35 36 37 38

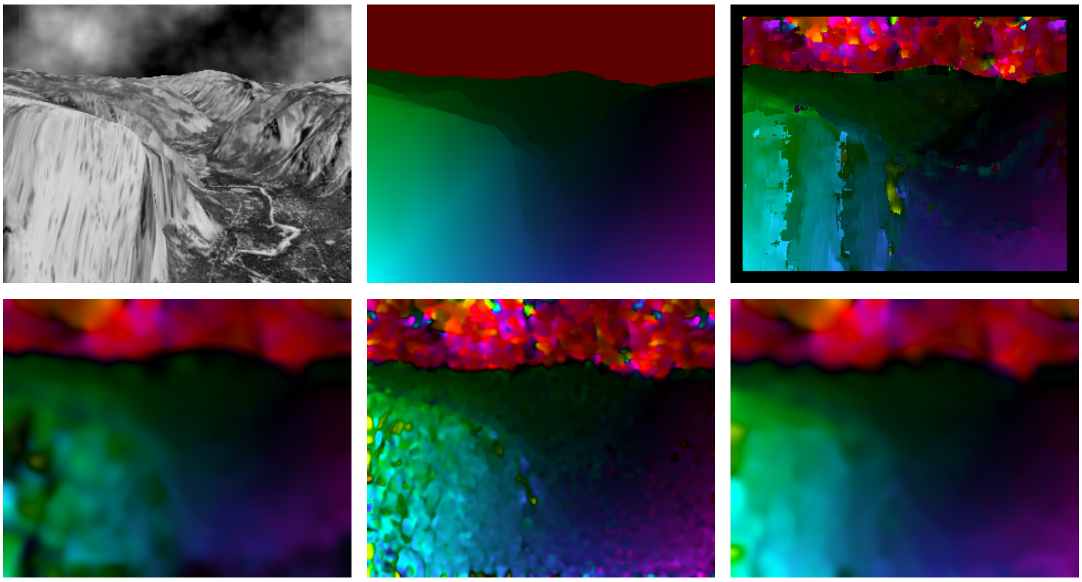
The Method of Lucas and Kanade (5)

Advantages and Shortcomings

- Advantages
  - similar to block matching, easy to implement, average performance
  - high efficiency due to direct minimisation (no exhaustive search $\rightarrow$ complexity $O(n)$ with small constant)
  - intrinsic sub-pixel precision by continuous formulation
  - robust under noise due to neighbourhood information
- Drawbacks
  - restriction to small displacements due to the linearisation of the BCCE (large displacement variants are possible but significantly more complex)
  - motion discontinuities not preserved (as block matching methods)
  - yields non-dense flow fields that require post-processing interpolation steps
  - neighbourhood not invariant under non-translational motion (e.g. rotations)

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 27 28 29 30 31 32 33 34 35 36 37 38

Results for the Method of Lucas and Kanade



Results for the Yosemite Sequence with clouds (L. Quam). **(a) Upper Left:** Frame 8. **(b) Upper Center:** Ground truth. **(c) Upper Right:** SDD + Subpixel ($m = 4$, $d = 7$). **(d) Lower Left:** Lucas/Kanade ($\sigma = 0$, $\rho = 6.3$). **(d) Lower Center:** Lucas/Kanade ($\sigma = 1.4$, $\rho = 2.0$). **(d) Lower Right:** Lucas/Kanade ($\sigma = 1.4$, $\rho = 6.3$).

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 27 28 29 30 31 32 33 34 35 36 37 38

Comparison in Terms of the Average Angular Error (AAE)

- Qualitative Evaluation for the Yosemite Sequence with Clouds

Technique	AAE
Basic Approach	72.89°
Zero Flow Field	55.22°
Block Matching (SSD)	24.44°
Block Matching (SAD)	24.40°
Normalised Cross Correlation (NCC)	21.84°
Block Matching + Subpixel (SSD)	21.46°
Lucas/Kanade	**16.28°**
Lucas/Kanade + Presmoothing	**8.79°**

- Flow fields significantly more accurate than the one of block matching methods due to sub-pixel precision and better treatment of boundaries.

The Spatiotemporal Variant of Lucas and Kanade

How Can We Improve Our Results Even Further?

- *Idea:* Use multiple frames and assume flow to be also constant over time
- The spatiotemporal variant of the LK-method minimises the local energy

$$E(u(x_0, y_0, t_0), v(x_0, y_0, t_0)) = \int_{\mathcal{N}_\rho(x_0, y_0, t_0)} \Big(f_x(x,y,t)\, u + f_y(x,y,t)\, v + f_t(x,y,t)\Big)^2 \, dx\, dy\, dt$$

  - $\mathcal{N}_\rho(x_0, y_0, \mathbf{t}_0)$: **spatiotemporal** neighbourhood window of size ρ
  - same minimisation procedure as in the spatial case
  - soft windowing realised by spatiotemporal Gaussian
- Advantage over the spatial variant: even higher robustness under noise

Comparison in Terms of the Average Angular Error (AAE)

- Qualitative Evaluation for the Yosemite Sequence with Clouds

Technique	**AAE**
Basic Approach	72.89°
Zero Flow Field	55.22°
Block Matching (SSD)	24.44°
Block Matching (SAD)	24.40°
Normalised Cross Correlation (NCC)	21.84°
Block Matching + Subpixel (SSD)	21.46°
Lucas/Kanade (2-D)	16.28°
Lucas/Kanade + Presmoothing (2-D)	8.79°
Lucas/Kanade + Presmoothing (3-D)	**7.69**°

- Accuracy of the results improves if more than two frames are used for the estimation. Are there alternatives to a local least squares fit?

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 27 28 29 30 31 32 33 34 35 36 37 38

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 27 28 29 30 31 32 33 34 35 36 37 38

The Method of Bigün et al.

How Can One Overcome the Aperture Problem?

- *Idea:* Apart from only considering spatiotemporal information also optimise the flow component in **temporal direction**
- The method of Bigün et al. minimises the local energy
  (Bigün/Granlund/Wiklund 1991)

$$E((\tilde{u},\tilde{v},\tilde{w})^\top) = \int_{\mathcal{N}_\rho(x_0,y_0,t_0)} \Big(f_x(x,y,t)\,\tilde{u} + f_y(x,y,t)\,\tilde{v} + f_t(x,y,t)\,\tilde{w}\Big)^2 \,dx\,dy\,dt$$

subject to the constraint $\sqrt{\tilde{u}^2+\tilde{v}^2+\tilde{w}^2} = 1$ to avoid the trivial solution.

  - final solution given by $u(x_0,y_0,t_0) = \frac{\tilde{u}}{\tilde{w}}$ and $v(x_0,y_0,t_0) = \frac{\tilde{v}}{\tilde{w}}$
  - $\mathcal{N}_\rho(x_0,y_0,t_0)$: spatiotemporal neighbourhood window of size ρ
  - can be understood as a **total least squares fit** w.r.t. to u and v

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 27 28 29 30 31 32 33 34 35 36 37 38

Formulation as Quadratic Form

- Using the compact notation $\tilde{\mathbf{w}} = (\tilde{u},\tilde{v},\tilde{w})^\top$ and $\nabla_3 f = (f_x,f_y,f_t)^\top$, the approach of Bigün et al. can be reformulated as **the quadratic form**

$$\begin{aligned} E(\tilde{\mathbf{w}}) &= \int_{\mathcal{N}_\rho(x_0,y_0,t_0)} \Big(\nabla_3 f(x,y,t)^\top \tilde{\mathbf{w}}\Big)^2 \,dx\,dy\,dt \\ &= \tilde{\mathbf{w}}^\top \Bigg[\int_{\mathcal{N}_\rho(x_0,y_0,t_0)} \Big(\nabla_3 f(x,y,t)\nabla_3 f(x,y,t)^\top\Big)\,dx\,dy\,dt\Bigg]\,\tilde{\mathbf{w}} \\ &= \tilde{\mathbf{w}}^\top J_\rho(\nabla_3 f(x,y,t))\,\tilde{\mathbf{w}} \end{aligned}$$

where $J_\rho(\nabla_3 f(x,y,t))$ is a symmetric and positive semi-definite 3×3 matrix.

- Using the Gaussian K_ρ as window yields the **spatiotemporal structure tensor**:

$$J_\rho(\nabla_3 f(x,y,t)) = \begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) & K_\rho * (f_x f_t) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) & K_\rho * (f_y f_t) \\ K_\rho * (f_x f_t) & K_\rho * (f_y f_t) & K_\rho * (f_t^2) \end{pmatrix}$$

The Method of Bigün et al. (3)

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 27 28 29 30 31 32 33 34 35 36 37 38

Minimisation as Eigenvalue Problem

- The minimiser of a quadratic form $E(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x}$ subject to $|\mathbf{x}| = 1$ is given by the eigenvector associated to the smallest eigenvalue of A.
- Let us consider the following eigenvalue decomposition $J_\rho(\nabla_3 f(x, y, t))$ which represents the matrix A in our case:

$$J_\rho(\nabla_3 f(x,y,t)) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)^\top$$

- Let us assume w.l.o.g that $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Then the solution is given by the eigenvector $\mathbf{e}_3$, while the corresponding minimum is given by $E(\mathbf{e}_3) = \lambda_3$.
- Evidently a smallest eigenvalue $\lambda_3 = 0$ is desirable, since it indicates that the minimum is $E(\mathbf{e}_3) = 0$ and thus that all local constraints are jointly fulfilled by the minimiser $\mathbf{e}_3$.

The Method of Bigün et al. (4)

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 27 28 29 30 31 32 33 34 35 36 37 38

Interpretation of the Eigenvalues

- This time we can distinguish four eigenvalue scenarios:
  - **Rank 3 – variation in three directions → contradictive flow**
  - Rank 2 – variation in two direction → full flow
  - Rank 1 – variation in one direction → normal flow
  - Rank 0 – no local structure → no flow
- While in the case of rank-3- or rank-2-matrices a unique solution for u and v exists, there are once more infinitely many solutions if two or more eigenvalues are zero, i.e. if the rank of the matrix comes down to 1 or 0.
- However, since the BCCE assumes that along the motion trajectory the grey value of a pixel remains constant (no variation), a rank of 3, i.e. $E(\mathbf{e}_3) = \lambda_3 \gg 0$, is not desirable. It indicates contradictive information due to noise or occlusions.
- For the method of Lucas and Kanade this case does not exist explicitly. However, one can also check for this method if the minimum of the least square fit is zero or not in ordert to detect problems.

The Method of Bigün et al. (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 27 28 29 30 31 32 33 34 35 36 37 38

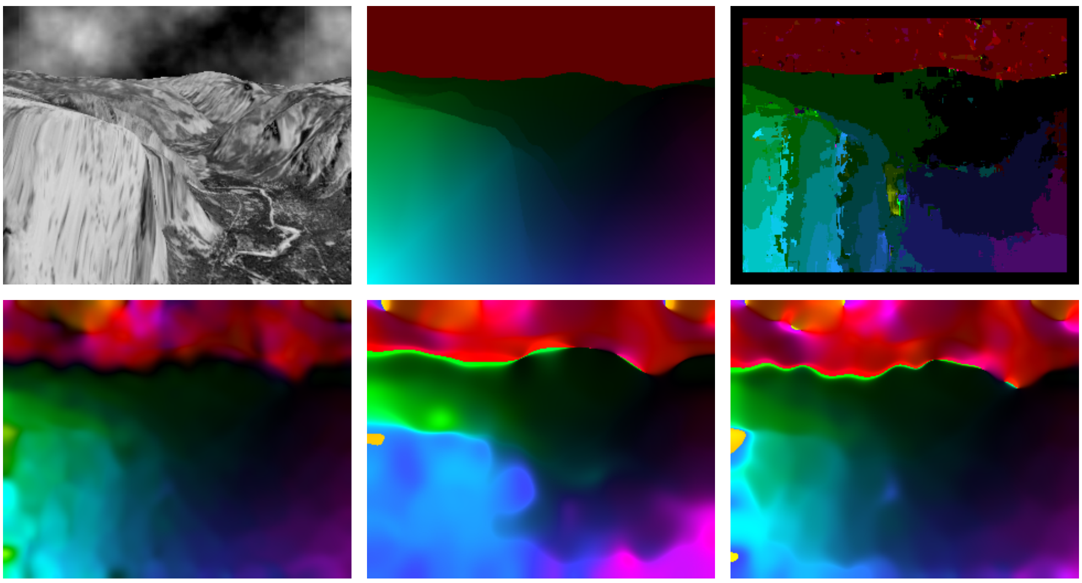
Advantages and Shortcomings

- Advantages
  - similar advantages as the method of Lucas and Kanade
  - optimises also the temporal optic flow component via a total least squares fit
  - can be more robust with respect to outliers and noise
- Drawbacks
  - similar drawbacks as the method of Lucas and Kanade
  - eigenvalue decomposition for 3×3 matrices computationally more demanding than solution of a 2×2 equation system
  - in order to distinguish the different eigenvalue scenarios, at least three thresholds are required (how to chose them appropriately?)

The Method of Bigün et al. (6)

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 27 28 29 30 31 32 33 34 35 36 37 38

Results for the Method of Bigün et al.



Results for the Yosemite Sequence with clouds (L. Quam). **(a) Upper Left:** Frame 8. **(b) Upper Center:** Ground truth. **(c) Upper Right:** NCC ($m = 4$, $d = 7$) **(d) Lower Left:** Lucas/Kanade 2-D ($\sigma = 1.4$, $\rho = 6.3$). **(d) Lower Center:** Bigün et al. 2-D ($\sigma = 0.0$, $\rho = 10.2$). **(d) Lower Right:** Bigün et al. 2-D ($\sigma = 1.6$, $\rho = 8.4$).

Comparison in Terms of the Average Angular Error (AAE)

- Qualitative Evaluation for the Yosemite Sequence with Clouds

Technique	AAE
Block Matching (SSD)	24.44°
Block Matching (SAD)	24.40°
Normalised Cross Correlation (NCC)	21.84°
Block Matching + Subpixel (SSD)	21.46°
Lucas/Kanade (2-D)	16.28°
Bigün et al. + Presmoothing (2-D)	**10.60°**
Bigün et al. + Presmoothing (3-D)	**9.15°**
Lucas/Kanade + Presmoothing (2-D)	8.79°
Lucas/Kanade + Presmoothing (3-D)	7.69°

- In some cases the method of Bigün is more precise than the method of Lucas and Kanade. In general, the difference in terms of quality is rather small.

Least Squares vs. Total Least Squares

Comparison of the Formulation of Lucas/Kanade and Bigün

- **Least squares fit:** Using $\mathbf{w} = (u, v, 1)^\top$ the method of Lucas/Kanade reads

$$E(\mathbf{w}) = \int_{\mathcal{N}_\rho(x_0,y_0,t_0)} \left(\nabla_3 f(x,y,t)^\top \mathbf{w}\right)^2 dx\,dy\,dt = \mathbf{w}^\top J_\rho(\nabla_3 f(x,y,t))\,\mathbf{w}\ .$$

- **Total least squares fit:** Using $\tilde{\mathbf{w}} = (\tilde{u}, \tilde{v}, \tilde{w})^\top$ the method of Bigün et al. reads

$$E(\tilde{\mathbf{w}}) = \int_{\mathcal{N}_\rho(x_0,y_0,t_0)} \left(\nabla_3 f(x,y,t)^\top \tilde{\mathbf{w}}\right)^2 dx\,dy\,dt = \tilde{\mathbf{w}}^\top J_\rho(\nabla_3 f(x,y,t))\,\tilde{\mathbf{w}}$$

subject to $|\tilde{\mathbf{w}}| = 1$. Solution is given by $u = \frac{\tilde{u}}{\tilde{w}}$ and $v = \frac{\tilde{v}}{\tilde{w}}$.

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 27 28 29 30 31 32 33 34 35 36 37 38

Least Squares vs. Total Least Squares (2)

Comparison of the Minimisation of Lucas/Kanade and Bigün

- **Least squares fit:** The method of Lucas/Kanade requires to solve the 2×2 linear system of equations

$$\begin{pmatrix} J_{\rho,11} & J_{\rho,12} \\ J_{\rho,12} & J_{\rho,22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -J_{\rho,13} \\ -J_{\rho,23} \end{pmatrix} .$$

- **Total least squares fit:** The method of Bigün et al. requires to determine the eigenvector that corresponds to the smallest eigenvalue of the 3×3 matrix

$$J_\rho = \begin{pmatrix} J_{\rho,11} & J_{\rho,12} & J_{\rho,13} \\ J_{\rho,12} & J_{\rho,22} & J_{\rho,23} \\ J_{\rho,13} & J_{\rho,23} & J_{\rho,33} \end{pmatrix} .$$

- In both cases the entries of the (spatiotemporal) structure tensor are required.

Parameterisation Models (1)

Parameterisation Models

Affine Parameterisation

- *Idea:* Assume the flow to be locally affine instead of locally constant
- Uses the following affine parameterisation model

$$\mathbf{w} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} ax+by+c \\ dx+ey+f \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} x & y & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x & y & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{M_{\text{affine}}} \underbrace{\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ 1 \end{pmatrix}}_{\mathbf{p}_{\text{affine}}}$$

- Optic flow constraint can be rewritten as

$$\nabla_3 f^\top \mathbf{w} = \nabla_3 f^\top (M_{\text{affine}}\, \mathbf{p}_{\text{affine}}) = (M_{\text{affine}}^\top \nabla_3 f)^\top \mathbf{p}_{\text{affine}} = \mathbf{r}_{\text{affine}}^\top\, \mathbf{p}_{\text{affine}} .$$

Parameterisation Models (2)

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 27 28 29 30 31 32 33 34 35 36 37 38

The Affine Lucas/Kanade Variant

- Plugging the affine parameterisation into the Lucas/Kanade method yields

$$E(\mathbf{p}_{\text{affine}}) \;=\; \int\limits_{\mathcal{N}_\rho(x_0,y_0,t_0)} \left(\mathbf{r}_{\text{affine}}(x,y,t)^\top \, \mathbf{p}_{\text{affine}}\right)^2 dx\,dy\,dt\;.$$

- Since we assume that the flow is affine, i.e. that $\mathbf{p}_{\text{affine}}$ is constant within the neighbourhood, we can reformulate the approach as

$$\begin{aligned} E(\mathbf{p}_{\text{affine}}) &= \mathbf{p}_{\text{affine}}^\top \Big[\int\limits_{\mathcal{N}_\rho(x_0,y_0,t_0)} \left(\mathbf{r}_{\text{affine}}(x,y,t)\; \mathbf{r}_{\text{affine}}(x,y,t)^\top\right) dx\,dy\,dt \Big]\; \mathbf{p}_{\text{affine}} \\ &= \mathbf{p}_{\text{affine}}^\top \; J_\rho(\mathbf{r}_{\text{affine}}(x,y,t))\; \mathbf{p}_{\text{affine}} \end{aligned}$$

where $J_\rho(\mathbf{r}_{\text{affine}}(x,y,t))$ is a symmetric and positiv semi-definite 7×7 matrix.

Parameterisation Models (3)

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 27 28 29 30 31 32 33 34 35 36 37 38

The Affine Lucas/Kanade Variant

- The minimum requires the derivatives w.r.t. all 6 parameters to be zero

$$\begin{aligned} 0 &\stackrel{!}{=} \frac{\partial E}{\partial a} \\ &\;\vdots \\ 0 &\stackrel{!}{=} \frac{\partial E}{\partial f} \end{aligned}$$

- This yields a 6×6 linear system of equations that has to be solved

$$\begin{pmatrix} J_{\rho,11} & J_{\rho,12} & J_{\rho,13} & J_{\rho,14} & J_{\rho,15} & J_{\rho,16} \\ J_{\rho,12} & J_{\rho,22} & J_{\rho,23} & J_{\rho,24} & J_{\rho,25} & J_{\rho,26} \\ J_{\rho,13} & J_{\rho,23} & J_{\rho,33} & J_{\rho,34} & J_{\rho,35} & J_{\rho,36} \\ J_{\rho,14} & J_{\rho,24} & J_{\rho,34} & J_{\rho,44} & J_{\rho,45} & J_{\rho,46} \\ J_{\rho,15} & J_{\rho,25} & J_{\rho,35} & J_{\rho,45} & J_{\rho,55} & J_{\rho,56} \\ J_{\rho,16} & J_{\rho,26} & J_{\rho,36} & J_{\rho,46} & J_{\rho,56} & J_{\rho,66} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} -J_{\rho,17} \\ -J_{\rho,27} \\ -J_{\rho,37} \\ -J_{\rho,47} \\ -J_{\rho,57} \\ -J_{\rho,67} \end{pmatrix} .$$

Parameterisation Models (4)

Rigid Body Parameterisation

- *Idea:* Assume the flow to be a projection of a 3-D rigid body motion given by the translation $\mathbf{t} = (t_1, t_2, t_3)$ and the rotation $\omega = (\omega_1, \omega_2, \omega_3)$
- Uses the following rigid body model (with unknown depth z)

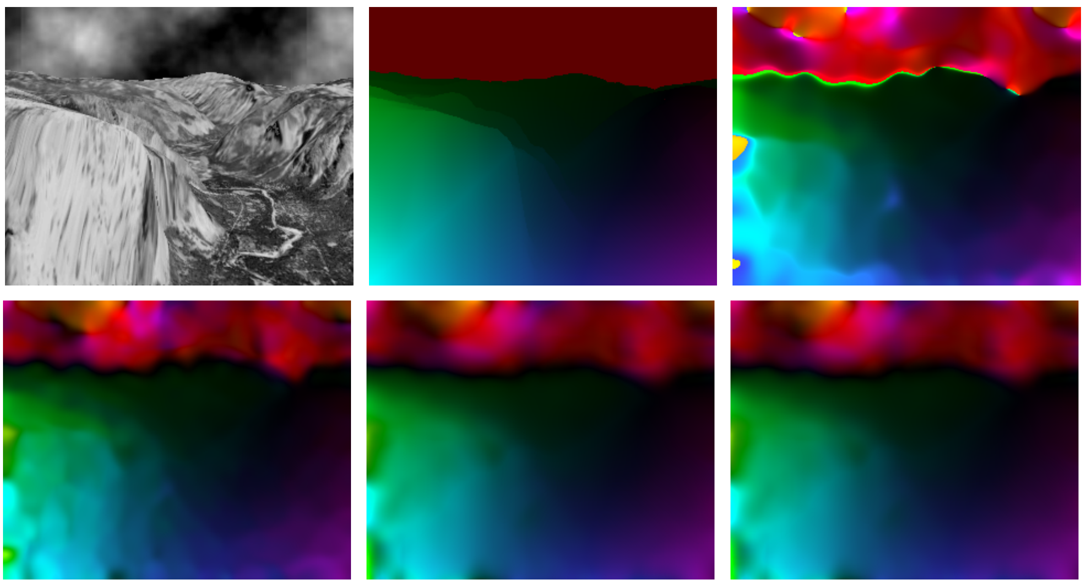
$$\mathbf{w} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{z}(ft_1 - xt_3) - \frac{xy}{f}\,\omega_1 + \frac{f^2+x^2}{f}\,\omega_2 - y\,\omega_3 \\ \frac{1}{z}(ft_2 - yt_3) - \frac{f^2+y^2}{f}\,\omega_1 + \frac{xy}{f}\,\omega_2 - x\,\omega_3 \\ 1 \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} f & 0 & -x & -\frac{xy}{f} & \frac{f^2+x^2}{f} & -y & 0 \\ 0 & f & -y & -\frac{f^2+y^2}{f} & \frac{xy}{f} & -x & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{M_{\text{rigid}}} \underbrace{\begin{pmatrix} \frac{t_1}{z} \\ \frac{t_2}{z} \\ \frac{t_3}{z} \\ \omega_1 \\ \omega_2 \\ \omega_3 \\ 1 \end{pmatrix}}_{\mathbf{p}_{\text{rigid}}}$$

- Requires knowledge on the focal length f of the camera

Parameterisation Models (5)

Results for the Affine and Rigid Body Lucas/Kanade Variant



Results for the Yosemite Sequence with clouds (L. Quam). **(a) Upper Left:** Frame 8. **(b) Upper Center:** Ground truth. **(c) Upper Right:** Bigün et al. 2-D ($\sigma = 1.6$, $\rho = 8.4$) **(d) Lower Left:** Lucas/Kanade 2-D ($\sigma = 1.4$, $\rho = 6.3$). **(d) Lower Center:** Affine Lucas/Kanade 2-D ($\sigma = 1.4$, $\rho = 6.3$). **(d) Lower Right:** Rigid Body Lucas/Kanade 2-D ($\sigma = 1.4$, $\rho = 6.3$).

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 27 28 29 30 31 32 33 34 35 36 37 38

Comparison in Terms of the Average Angular Error (AAE)

- Qualitative Evaluation for the Yosemite Sequence with Clouds

Technique	AAE
Normalised Cross Correlation (NCC)	21.84°
Block Matching + Subpixel (SSD)	21.46°
Lucas/Kanade (2-D)	16.28°
Bigün et al. + Presmoothing (2-D)	10.60°
Bigün et al. + Presmoothing (3-D)	9.15°
Lucas/Kanade + Presmoothing (2-D)	8.79°
Lucas/Kanade + Presmoothing (3-D)	7.69°
Affine Lucas/Kanade + Presmoothing (2-D)	**7.53°**
Rigid Body Lucas/Kanade + Presmoothing (2-D)	**7.52°**

- The affine and the rigid body parameterisation yield slight improvements.

Summary

- Block matching results are not convincing in 2-D
- Continuous modelling yields real sub-pixel precision and thus better results
- The methods of Lucas and Kanade and Bigün *et al.*
  - make use of least squares and total least squares fitting approaches
  - can be applied in a spatial and in a spatiotemporal setting
  - allow for a direct minimisation and are thus computationally efficient
  - require the solution of a 2×2 system or a 3×3 eigenvalue decomposition
  - can be implemented by using finite difference approximations
- Alternative parameterisations of the optic flow may improve the results
  - affine parameterisation
  - rigid-body parameterisation

Summary (2)

Literature

- B. Lucas, T. Kanade:
  An iterative image registration technique with an application to stereo vision.
  In *Proc. 7th International Joint Conference on Artificial Intelligence*, pp. 674–679, 1982.
  (description of the method of Lucas and Kanade)
- S. Baker, I. Matthews:
  Lucas–Kanade 20 years on: A unifying framework.
  In *International Journal of Computer Vision*, Vol. 56, No. 3, pp. 221–255, 2004.
  (detailed description of all ingredients of the Lucas and Kanade algorithm)
- J. Bigün, G. H. Granlund, J. Wiklund:
  Multidimensional orientation estimation with applications to texture analysis and optical flow.
  In *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 8, pp. 775–790, 1991.
  (description of the method of Bigün et al.)

Assignment 2

Assignment 2

Theoretical Exercise 1 (Sub-Pixel Refinement I)

Let the three cost values $f(0) = c_0$, $f(-1) = c_{-1}$ and $f(1) = c_{+1}$ for the best match and its left and right neighbour be given. Show that the fitted parabola $f(x) = ax^2 + bx + c$ for the SSD model reads:

$$\frac{c_{+1} - 2c_0 + c_{-1}}{2}\, x^2 + \frac{c_{+1} - c_{-1}}{2}\, x + c_0 = 0$$

(*Hint: Set up a linear equation system and solve for a, b and c.*)

Theoretical Exercise 2 (Sub-Pixel Refinement II)

How does the fitted absolute value function $f(x) = a|x + b| + c$ for the SAD model look like. You may solve the problem geometrically.

(*Hint: One knows that $c_{-1} \geq c_0$ and $c_{+1} \geq c_0$ and thus that $-1 \leq b \leq 1$. The solution is given by*

$$f(x) = \begin{cases} (c_{+1} - c_0)\left|x + \frac{c_{+1} - c_{-1}}{2(c_{+1} - c_0)}\right| + \frac{c_{-1} + 2c_0 - c_{+1}}{2} & \text{for} \quad c_{+1} \geq c_{-1} \\ (c_{-1} - c_0)\left|x + \frac{c_{-1} - c_{+1}}{2(c_{-1} - c_0)}\right| + \frac{c_{+1} + 2c_0 - c_{-1}}{2} & \text{for} \quad c_{+1} < c_{-1} \end{cases}$$

)