Robot Vision

TTK4255

Lecture 10 – Motion Estimation and Tracking

Annette Stahl

(Annette.Stahl@ntnu.no)

Department of Engineering Cybernetics – ITK

NTNU, Trondheim

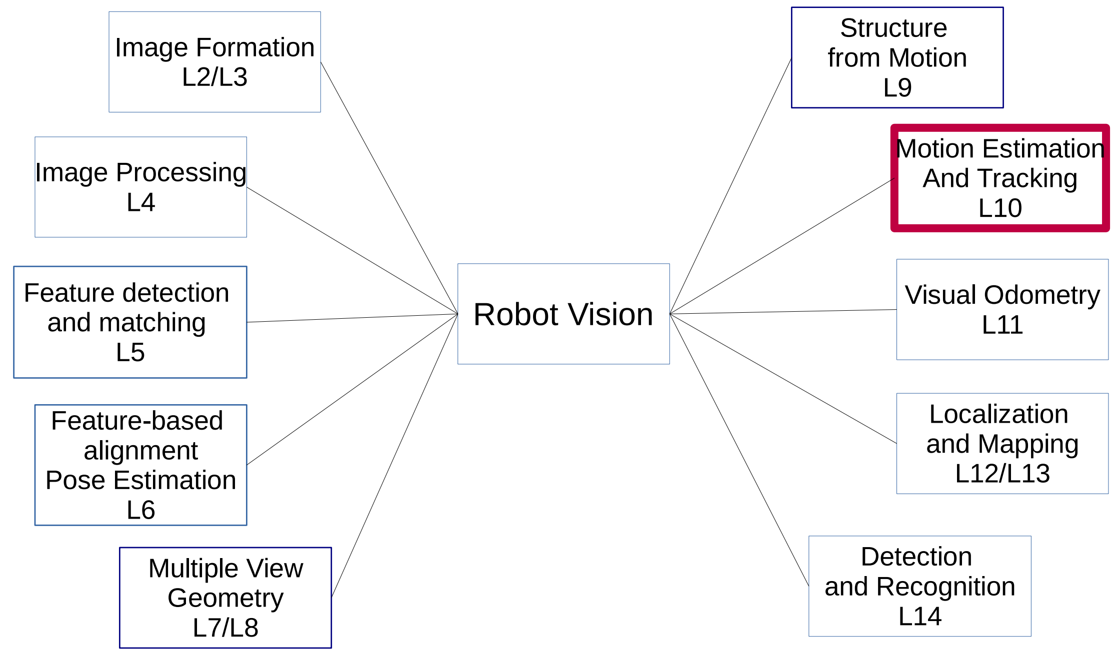
Spring Semester

09. March 2020

Lecture 10 – Motion Estimation

Annette Stahl (Annette.Stahl@ntnu.no)

Simen Haugo (Simen.Haugo@ntnu.no)



Outline of the tenth lecture:

- Variational methods
- Variational image smoothing
- Variational image motion estimation
- Euler-Lagrange Equation
- Gradient Decent
- Coarse-to-Fine Strategy
- Applications

Variational Methods

- Variational methods are a specific sub-class of optimization methods.
- They are popular because they allow to solve many problems in a mathematically transparent manner.
- Instead of implementing a heuristic sequence of processing steps as one clarifies beforehand what properties an optimal solution should have
- Variation methods are particularly popular for infinite dimensional problems and spatially continuous representation

Particular applications are

- Image denoising and image restoration
- Image segmentation
- Motion estimation and optical flow
- Spatially dense multiple view reconstruction
- Tracking

Advantages of Variational Methods

Variational methods have many advantages over heuristic muti-step approaches:

- A mathematical analysis of the considered cost function allows to make **statements on the existence and uniqueness of solutions**
- Approaches with multiple processing steps are difficult to modify. All steps rely on the input from a previous step. Exchanging one module by another typically requires to re-engineer the entire processing pipeline
- Variational methods make all **modeling assumptions transparent**, they are no hidden assumptions
- Variational methods typically have **fewer tuning parameters**. In addition, the effect of respective parameters is clear
- Variational methods are **easily fused** – one simply adds respective energies /cost functions

Example: Variational Image Smoothing

Let $f : \Omega \to \mathbb{R}$ be a gray value input image on the domain $\Omega \subset \mathbb{R}^2$

Assumption: the observed image arises by some true image corrupted by additive noise.

Aim: De-noise version of u of the input image f

The approximation u should fulfill two properties

- It shoud be as similar as possible to f
- It should be spatially smooth (noise-free)

Both to these criteria can be formulated by using a cost function of the form

$$E(u) = E_{data}(u, f) + E_{regularity}(u)$$

First first term measures the similarity of f and u.

The second term measures the smoothness of $u.$

Most variational approaches have the above form. They merely differ in the specific form of the **data (similarity) term** and the **regularity (or smoothness term)**.

Variational Image Smoothing

For de-noising a gray value image $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$, specific examples of data and smoothness term are:

$$E_{data} = \int_\Omega (u(x,y) - f(x,y))^2 dxdy$$

And

$$E_{regularity} = \int_\Omega |\nabla u(x,y)|^2 dxdy$$

where $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)^\top$ denotes the spatial gradient.

Minimizing the weighted sum of data and smoothness term

$$E(u) = \int_\Omega (u(x,y) - f(x,y))^2 dxdy + \lambda \int_\Omega |\nabla u(x,y)|^2 dxdy, \quad \lambda > 0,$$

Leads to a smooth approximation $u : \Omega \to \mathbb{R}$ of the input image.

Such energies which assign a real value to a function are called a **functionals**. How does one minimize functionals where the argument is a function $u(x)$ (rather than a finite number of parameters)?

Minimization and Euler-Lagrange Equations

As a **necessary condition for minimizers of a functional** the associated **Euler-Lagrange equation** must hold. For a functional of the form

$$E(u) = \int L(u, u')dxdy$$

It is given by

$$\frac{dE}{du} = \frac{\partial L}{\partial u} - \frac{d}{dx}\frac{\partial L}{\partial u'} = 0$$

The central idea of variational methods is therefore to **determine solutions of the Euler-Lagrange equation** of a given functional. For general **non-convex functionals** this is a difficult problem.

Another solution is to start with an (appropriate) function $u_0(x)$ and to modify it step by step such that in each iteration the value of the functional is decreased. Such methods are called **descent methods**.

Gradient Descent

One specific descent method is called **gradient descent** or **steepest descent**. The key idea is to start from an initialization $u(x,t=0)$ and iteratively march in direction of the negative energy gradient.

For the class of functionals considered above, the gradient descent is given by the following **partial differential equations**:

$$\begin{cases} u(x,0) = u_0(x) \\ \frac{\partial u(x,t)}{\partial t} = -\frac{dE}{du} = -\frac{dL}{du} + \frac{\partial}{\partial x}\frac{dL}{du'} \end{cases}$$

Specifically for $L(u,u') = \frac{1}{2}(u(x) - f(x))^2 + \frac{\lambda}{2}|u'(x)|^2$ this means

$$\frac{\partial u}{\partial t} = (f - u) + \lambda u''$$

If the gradient descent evolution converges: $\frac{\partial u}{\partial t} = -\frac{dE}{du} = 0$ then we have found a solution for the Euler-Lagrange equation.

TV regularization and l2-norm

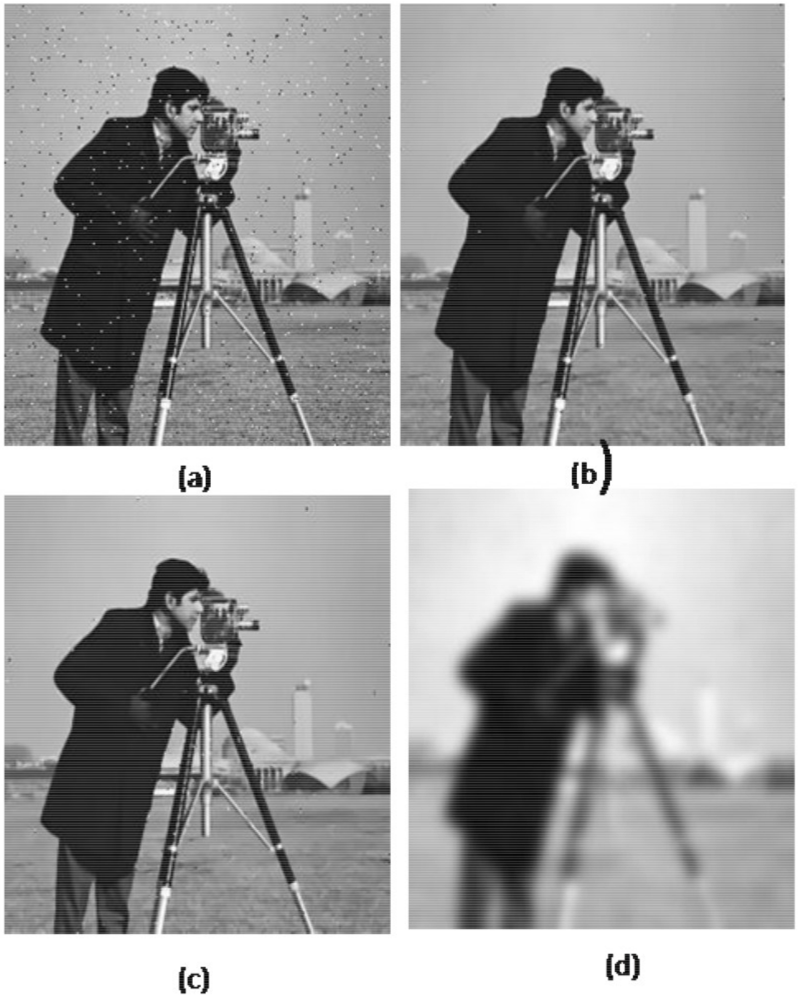


Fig 5. (a) noisy image (b)denoised image ROF model λ=0.003 (b) denoised image TV-L1 model λ=0.003 (d) denoised image Tikhonov model λ=0.003

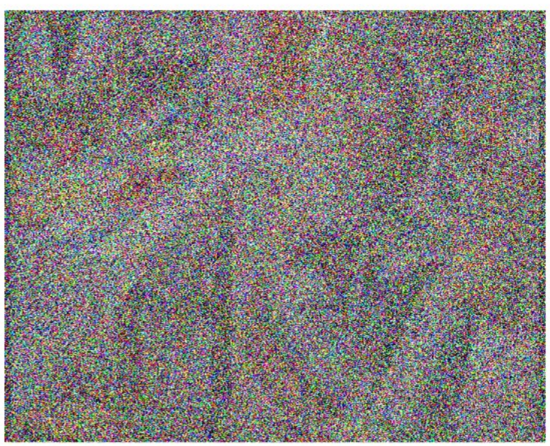
Vanjigounder, Kamalaveni & K.A., Narayanankutty & Veni, S.. (2016). Performance Comparison of Total Variation based Image Regularization Algorithms. International Journal on Advanced Science, Engineering and Information Technology.

TV - regularization

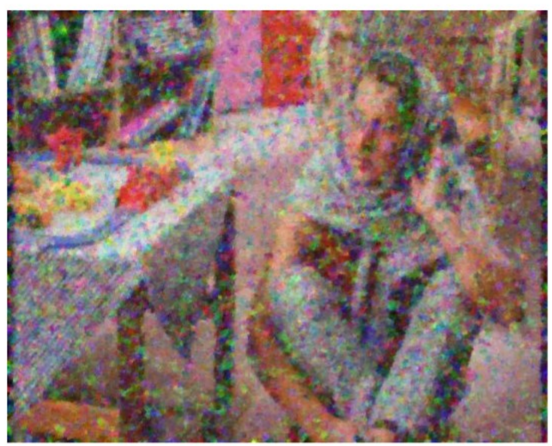
(a) Grayscale "Boats" image corrupted with 50% Salt & Pepper noise

(b) ℓ^1-TV denoising with regularization parameter $\lambda = 1.4$ (SNR: 9.01 dB, SSIM: 0.66)

(c) Adaptive ℓ^1-TV denoising (SNR: 15.81 dB, SSIM: 0.89)



(d) Color "Barbara" image corrupted with 80% Salt & Pepper noise



(e) ℓ^1-TV denoising with regularization parameter $\lambda = 1.5$ (SNR: 4.52 dB)

(f) Adaptive ℓ^1-TV denoising (SNR: 10.89 dB)

Rodríguez, Paul. Total Variation Regularization Algorithms for Images Corrupted with Different Noise Models: A Review. Journal of Electrical and Computer Engineering, Open Access, 2013.

Motion Representation

Image sequences or videos allow for an **analysis of dynamic processes**:

- transport/diffusion processes, biological growth
- traffic observation, (autonomous) robots

→ Determination of motion is useful for all **temporal changes** that can be recorded

t $\quad$ $t + \Delta t$ $\quad$ $t + 2\Delta t$

Image Sequence:

The observed displacements of objects between consecutive images is interpreted as the “motion“ of the objects.

Motion Estimation

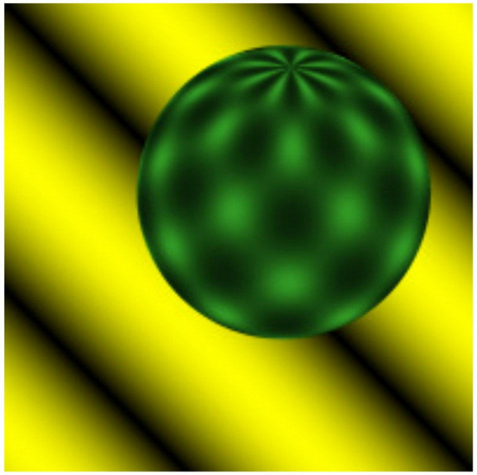
Central problem in Computer Vision

Estimation of motion fields from image sequences

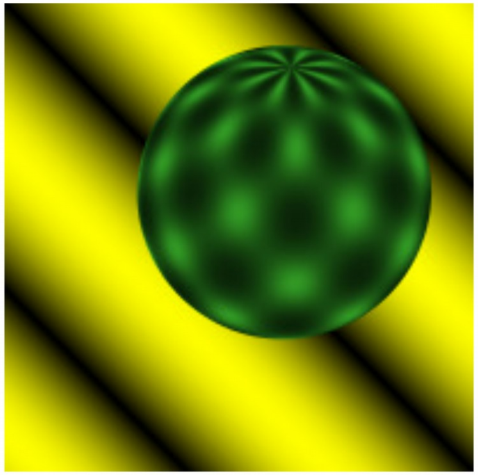
- **Increasing data amount:** Image sequence data captured from video capable cameras with higher frame rate, videos on the Internet → image sequence analysis has become increasingly important
- **Exploit temporal information:** Image sequences contain more information about the world surrounding us in the sense that structures can often be distinguished based on their temporal evolution → releases the sense of attention – compared against just still images
- **Camera Software:** some applications of motion estimation are integrated in camera softwere like panomrama generation from several images, video stabilization to remove camera motion sidturbences.
- **Ill-posed problem:** Mathematically the problem of motion estimation from images is ill-posed → the problem is not sufficiently specified to assure a unique solution (missing constraints with respect to the brightness consistency constraint)

Motion Estimation

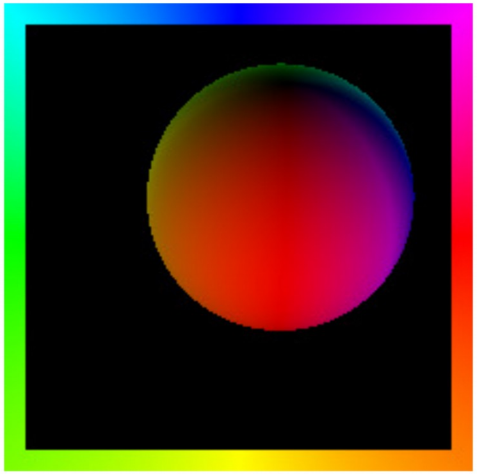
sphere – frame 5



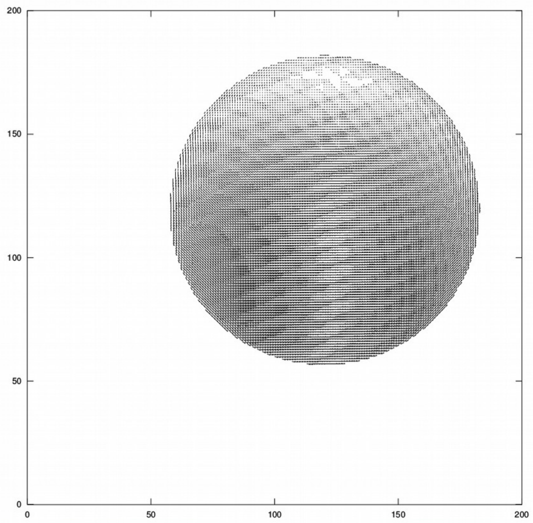
sphere – frame 6



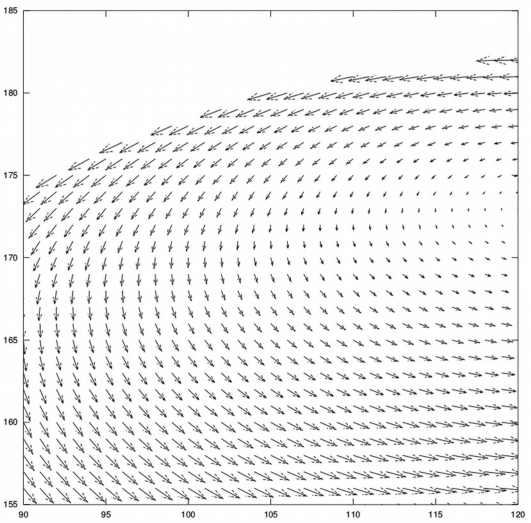
color representation of ground truth motion field



Ground truth motion



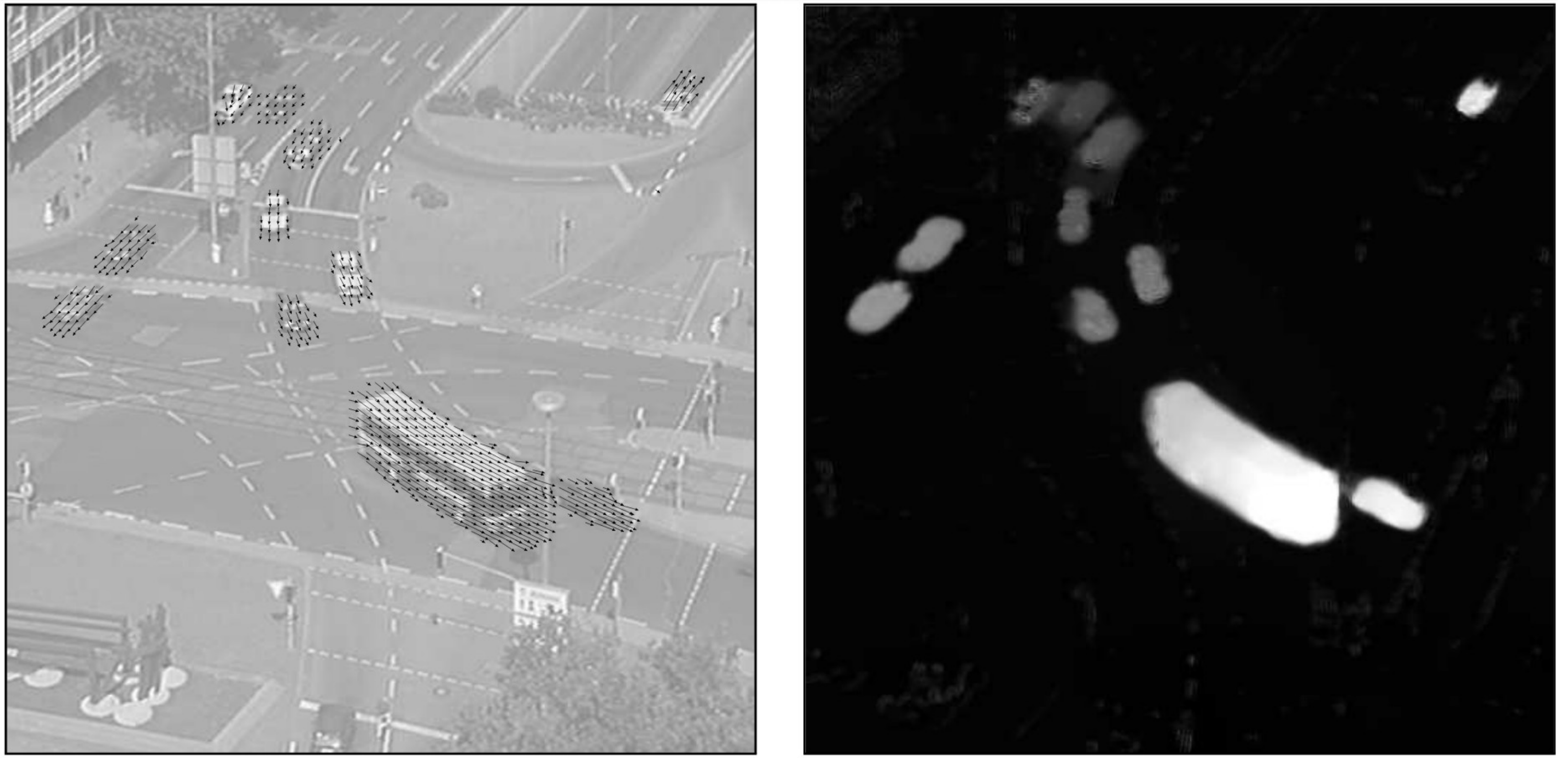
vector field (v_1, v_2) (zoom)



Motion estimation – Image Sequence

Application areas:

- Video compression
- 3D reconstruction (depth information, stereo)
- Segmentation



→ Aim: Approximate the motion field as good as possible !

Applications of Motion Estimation

- **Grouping and Segmentation:** Motion information allows to identify image regions as objects. This can also be done if semi-transparent structures overlap at a given location
- **Tracking:** Using motion information, objects can be tracked in a video sequence.
- **Depth estimation:** Motion information allows to infer the distance of respective objects from the camera. In principle one can recover the 3D geometry of the world from an image sequence
- **Diver assistance:** Motion allows to make predictions when an obstacle will be hit. → initiate evasion maneuvers or breaking
- **Video compression:** motion information allows to efficiently compress videos mpeg encoding
- **Mixing and dispersion:** motion information can give insights into mixing processes of fluids and dispersion of particles

Correspondence Problem

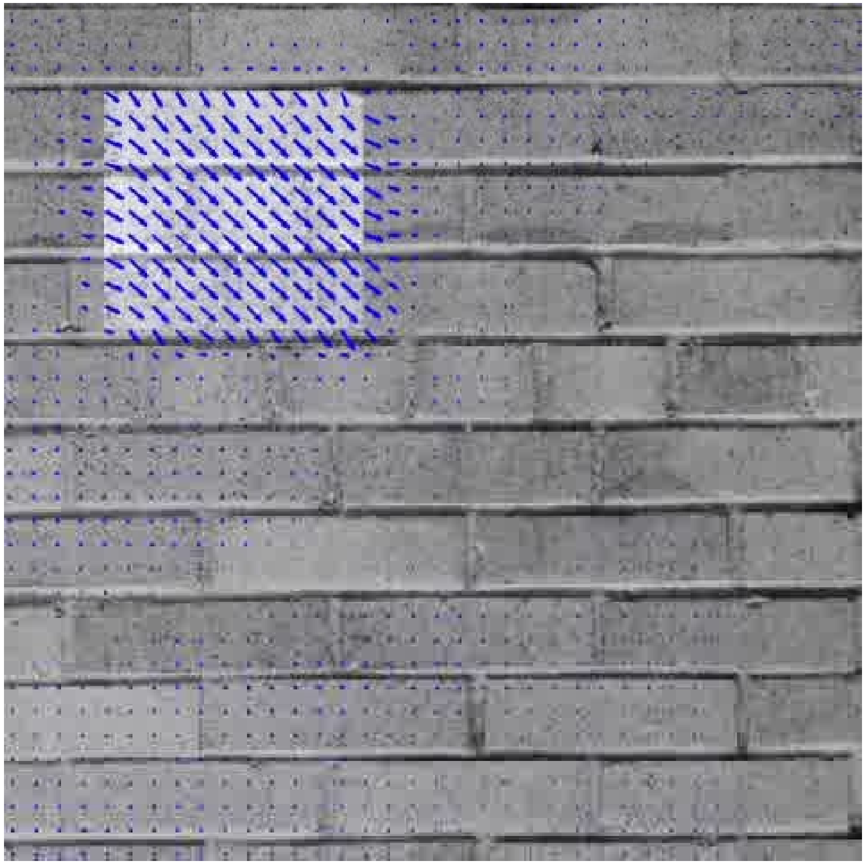
Key challenge in motion estimation:

solve the correspondence problem given two images. Determine for each point in either image the corresponding pixel location in the other image.

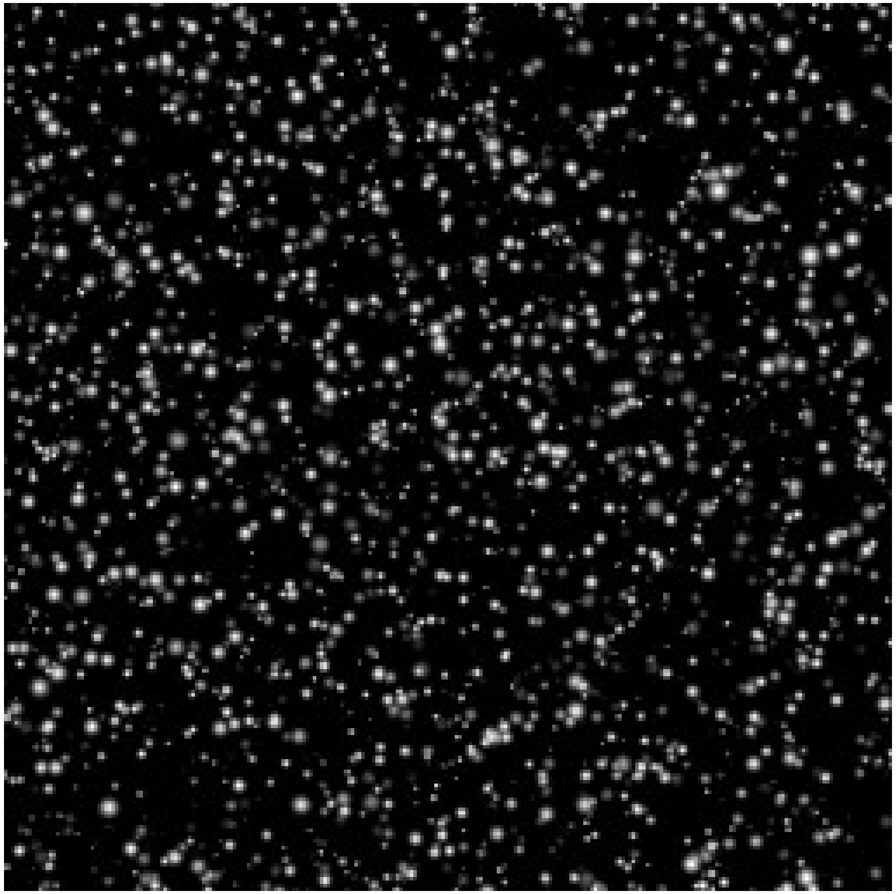
Related problems:

- **Disparity estimation from stereo images:** Determine a one-dimensional displacement for each pixel to determine the corresponding pixel in the other image.
- **Multimodal registration** (optical flow estimation in medial imaging): Given two medical images of an organ acquired with different sensors – CT (Computer Tomography) and MRI (magnet resonance imaging) – how to compute an optimal alignment of these images
- **Shape Matching:** Given two object shapes contours in 2D or surfaces in 3D determine a correspondence between pairs of points from either shape

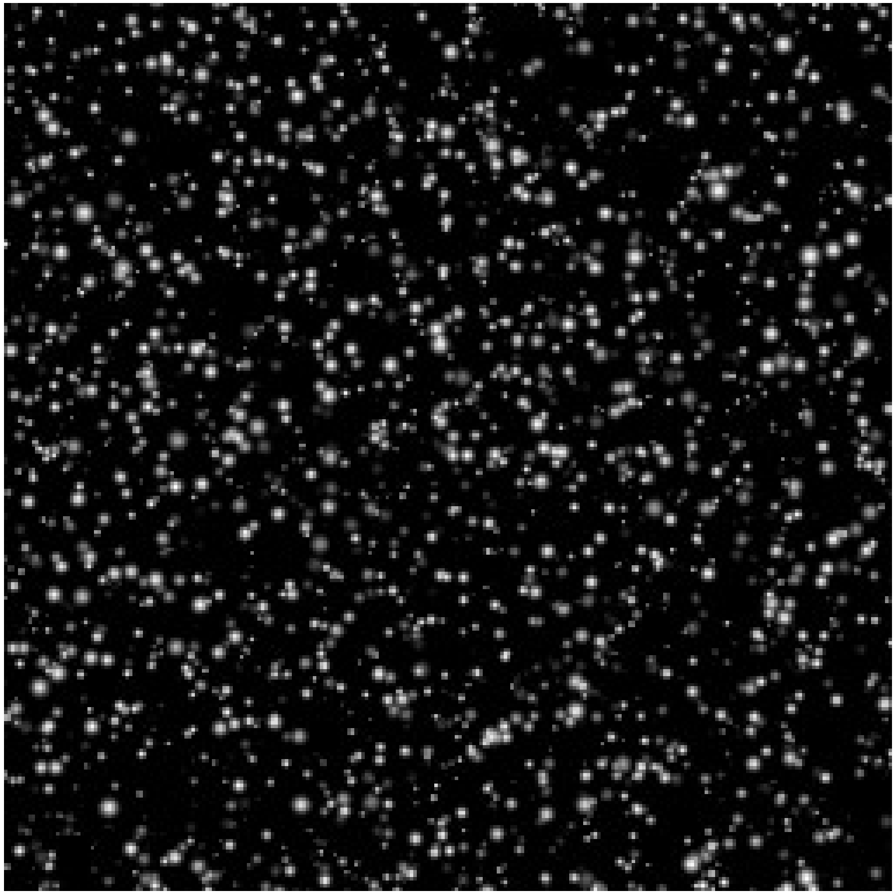
Correspondence Problem

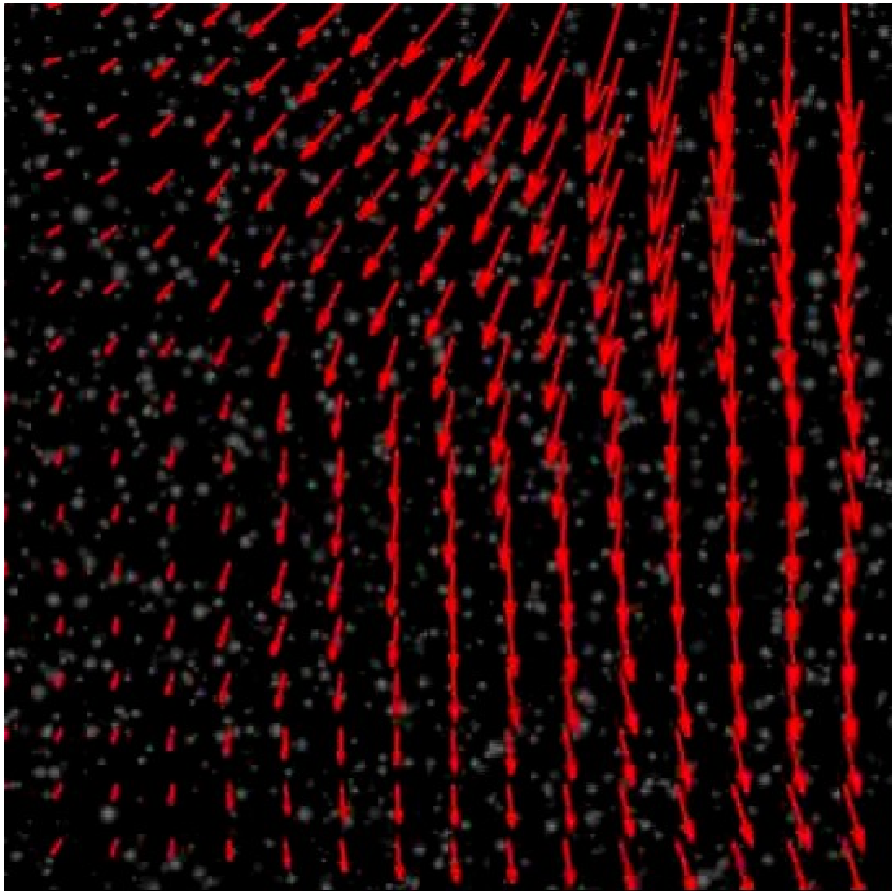


Correspondence Problem



Correspondence Problem





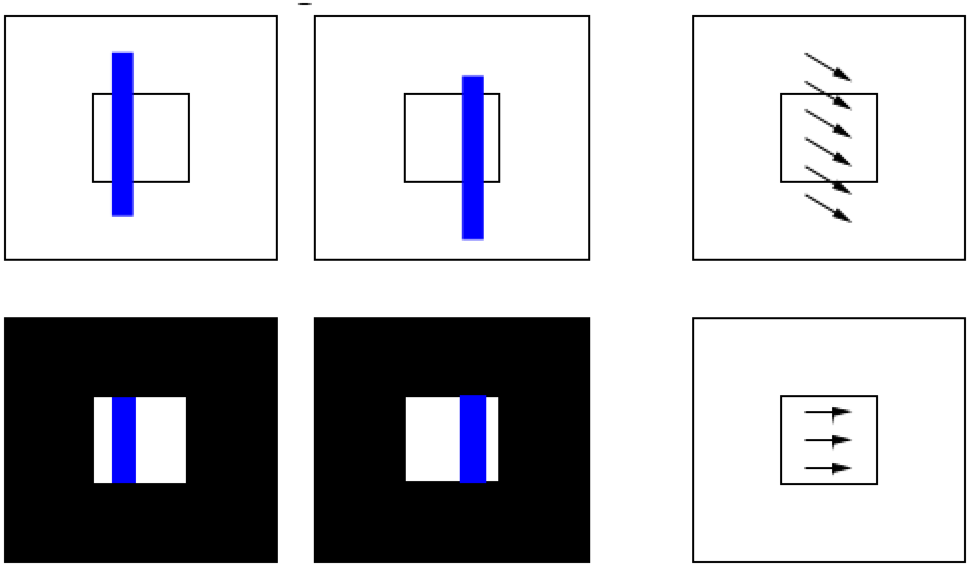
Aperture Problem

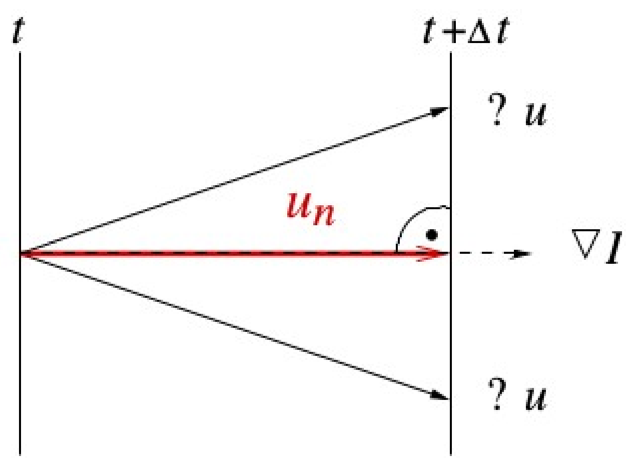
In general:

One cannot estimate the motion in direction of constant brightness (for example along an image edge). This limitation is referred to as the **aperture problem**.

For example:

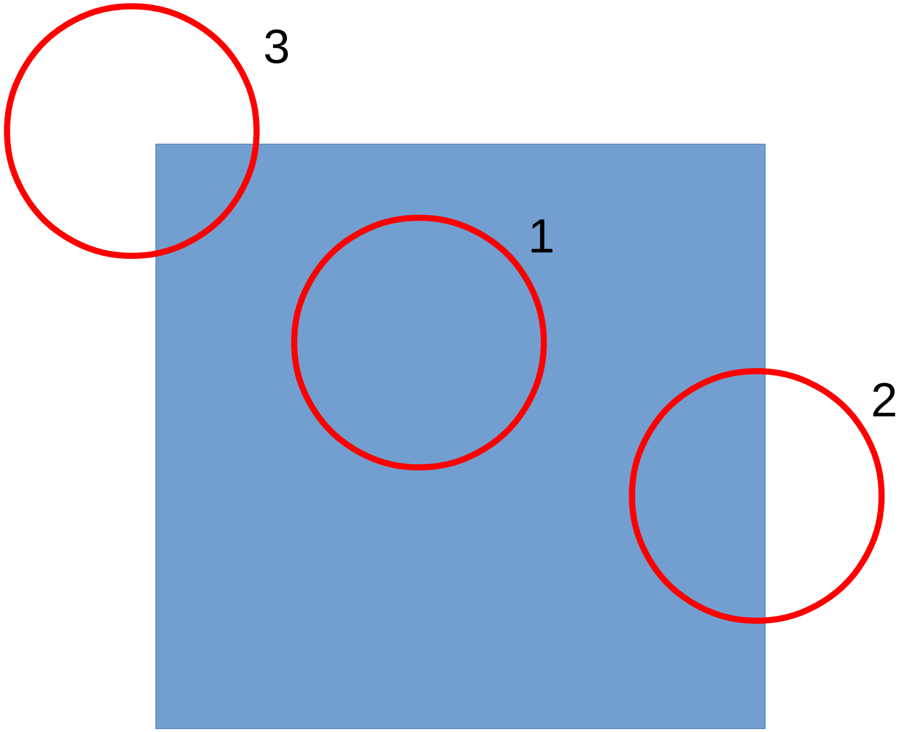
No matter how the vertical stripe pattern behind the mask is displaced we will only observe its horizontal motion.





Aperture Problem

Example: Consider three observers watching a local patch (red circle) of a moving blue box.



Observer 1 observes no motion

Observer 2 observes only horizontal motion

Observer 3 observes motion in both directions

Motion Field and Optical Flow

We **distinguish** between a motion field and the optical flow.

- **Motion Field:** The projection of the real 3D motion onto the image plane.
- **Optical Flow Field:** Apparent motion due to the gray value changes from one image to another image. (The optical flow field transforms the images into each other.)

→ the optical flow is only an approximation of the real motion!

Examples:

Not all visual changes are caused by motion:

- Reflections, illumination changes, shadows

Not all motions lead to visual changes:

- Moving homogeneous plane, rotating uniform colored sphere ,

Motion Analysis

Simple situations:

- Static scene, moving objects (non-moving camera, background fixed)

→ difference images might help to detect and localize motion (problem: homogeneous areas)

Difficult situations:

- Moving camera, non-static background, large displacements of the objects

→ matching techniques might help (find correspondences, match object features, match patches, graph matching, exploit knowledge about maximum velocity/small acceleration)

Medium difficult situations:

- Continuous small motions of camera and/or objects, nearly no illumination change

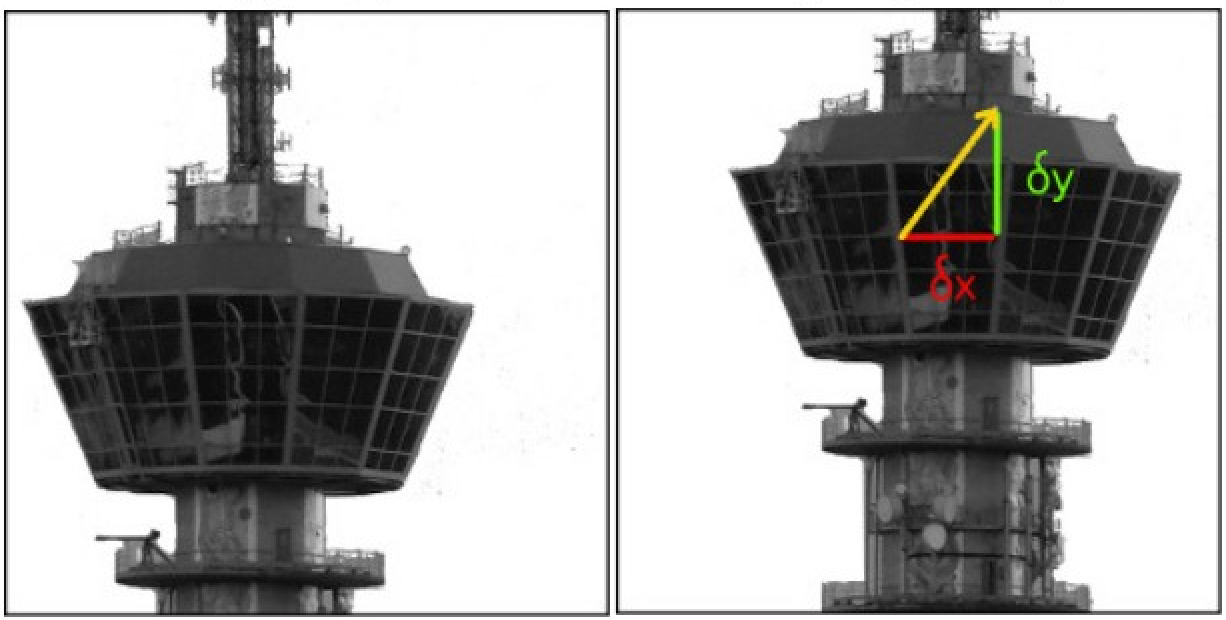
→ differential optical flow methods

Optical Flow Constraint

Assumption: The brightness of a patch that is shifted by $(\partial x, \partial y)$ in the time ∂t remains constant.

$$I(x, y, t) = I(x + \partial x, y + \partial y, t + \partial t)$$

Image $I(t)$ | Image $I(t + \delta t)$



at time t | at time $t + \delta t$

Derivation of the Optical Flow Constraint

Velocity is a differential quantity.

Idea: Estimation through the use of derivative measurements

Assumption of the gradient formulation of the optical flow problem:

Intensity conservation over time.

(Holds, if spacial and temporal sampling is sufficient fine!)

Meaning: Changes in image intensity due to translation of local image intensity and not to changes in illumination!

→ the total derivative with respect to time of the image intensity function should be zero at each position in the image and at every point in time!

Brightness Constancy Assumption

Given: An image sequence $I : \Omega \times [0, T] \to \mathbb{R}$ on the image plane $\Omega \subset \mathbb{R}^2$ and the time interval $[0, T]$

Aim: Compute a motion field $u : \Omega \times [0, T] \to \mathbb{R}^2$ which assigns to each point $(x, y) \in \Omega$ at each time $t \in [0, T]$ a motion vector $u(x, y, t) = [u_1, u_2]^\top$

Let $(x, y) : [0, T] \to \Omega$ denote the trajectory of an object point over time. The classical assumption in motion estimation states that the brightness of a moving point remains constant over time!

With writing the image intensity signal as continuous function of position and time $I(x(t), y(t), t)$ this means

$$I(x(t), y(t), t) = const. \quad \forall\, t \in [0, T]$$

We assume the brightness function to be differentiable, then we can deduce that the total time derivative must vanish:

$$\frac{dI(x(t), y(t), t)}{dt} = 0 \quad \forall t \in [0, T]$$

Optical Flow Constraint

This leads to the basic equation:

$$I(x, y, t) = I(x + \partial x, y + \partial y, t + \partial t)$$

If we perform the first order Taylor approximation of $I(x,y,t)$ and throw out nonlinear terms:

$$\begin{aligned} I(x, y, t) &\approx I(x, y, t) + \partial x \frac{\partial I}{\partial x} + \partial y \frac{\partial I}{\partial y} + \partial t \frac{\partial I}{\partial t} \\ &= I(x, y, t) + \partial x I_x + \partial y I_y + \partial t I_t \end{aligned}$$

→ subtracting $I(x,y,t)$ and dividing by ∂t leads to

$$\frac{\partial x}{\partial t} I_x + \frac{\partial y}{\partial t} I_y + I_t = 0$$

Optical Flow Constraint

$$I_x \frac{dx}{dt} + I_y \frac{dy}{dt} + I_t = 0 \quad \Longleftrightarrow \quad I_x \tilde{u} + I_y \tilde{v} + I_t = 0$$

Where the two components u and v of the optical flow are

$$\tilde{u} = \frac{dx}{dt} \quad \tilde{v} = \frac{dy}{dt} \quad \rightarrow \quad u = \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}$$

By using $\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$ the optical flow constraint can be stated as

$$\boxed{u^\top \cdot \nabla I = -I_t}$$

Flow components in normalized gradient direction $\frac{\nabla I}{|\nabla I|}$:

$$u_n = u^\top \cdot \frac{\nabla I}{|\nabla I|} = -\frac{I_t}{|\nabla I|}$$

Note: $u^\top$ is not uniquely determined → Aperture Problem

Optical Flow Constraint

The term $u(x, y, t) = \frac{dx}{dt}$ is nothing but the **velocity of the moving point** that we are looking for.

Thus the **assumption of brightness constancy and differentiability** lead to a relation between the desired velocity field $u(x, y, t)$ and the spatial and temporal image gradients:

$$u^\top \nabla I + I_t = 0$$

This equation is referred to the **differential brightness constancy constraint** or the **optical flow constraint (OFC).**

The optical flow constraint reflects the previously discussed aperture problem: It does not allow statements regarding motion along the level lines of constant intensity.

Aperture Problem

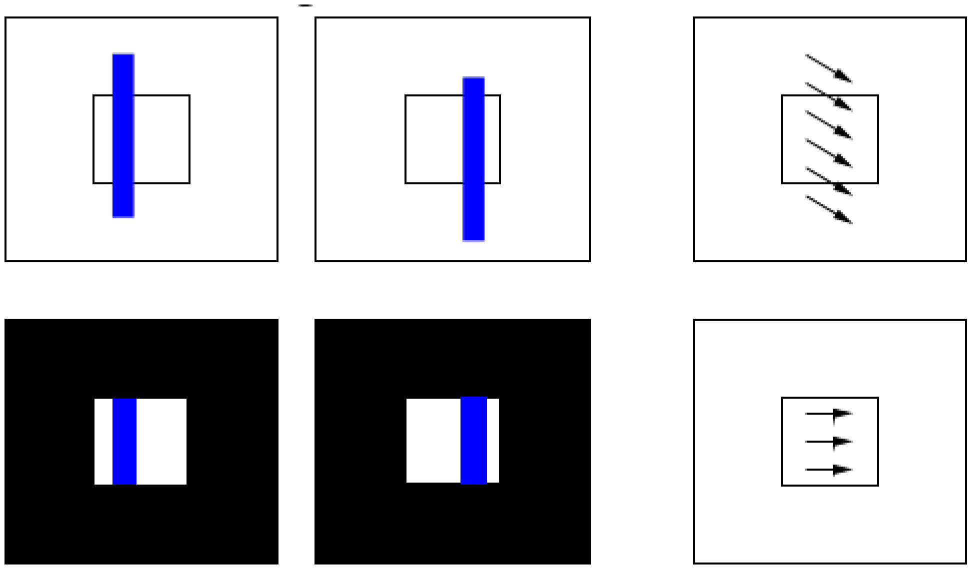
The Aperture Problem:

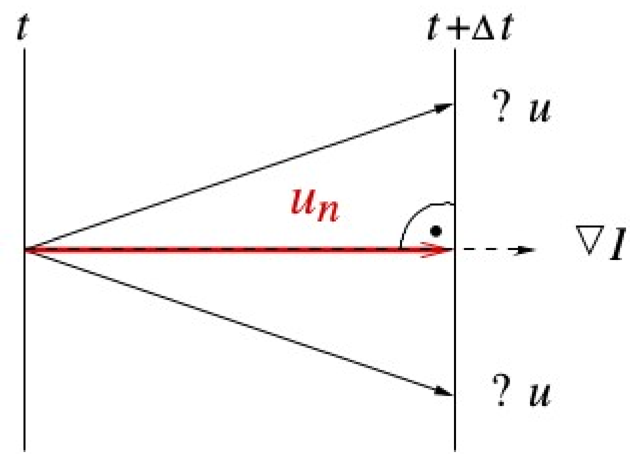
Only the optical flow component u_n in normalized gradient direction $\frac{\nabla I}{|\nabla I|}$ is determined $u_n = -\frac{I_t}{|\nabla I|}$

→ **ill-posed** (1 equation, 2 variables), requires additional constraints for unique flow estimate.
We can even see the non-uniqueness of the optical flow constraint equation.

→ Only the motion in the direction of the gradient ∇I is perceived.

→ The optical flow does not always correspond to the real motion field.





OFC and Aperture Problem

The aperture problem is reflected in the optical flow constraint because the constraint is invariant to changes in the motion field which are orthogonal to the local image gradient.

The central problem in motion estimation lies in the fact that the constraint coupling the velocity field $u(x,y,t)$ and the image gradients **cannot be directly solved for $\boldsymbol{u}$.**

More specifically, the flow constraint provides the projection $u_\perp$ of the velocity vector u onto the image gradient ∇I. Dividing the OFC by ∇I lead to

$$u_\perp \equiv \left(\frac{\nabla I}{|\nabla I|}\right)^\top \qquad u = -\frac{I_t}{\nabla I}$$

This component of the velocity normal to the level lines is called the **normal flow**. It is simply given by the negative ratio of temporal and spatial image gradient.

Additional Assumptions

The optical flow constraint is necessary but not sufficient to uniquely determine a motion field. It only specifies the normal component of the velocity field

In order to eliminate the additional degree of freedom we therefore need to make additional assumptions

Two pioneering approaches:

- **Lucas and Kanade** 1981 → **local** approach
  - B. D. Lucas and T. Kanade (1981), An iterative image registration technique with an application to stereo vision. Proceedings of Imaging Understanding Workshop, pages 121–130
  - cited ca. 15000 in 2020.
- **Horn and Schunck** 1981 → **global** approach
  - B.K.P. Horn and B.G. Schunck, "Determining optical flow." Artificial Intelligence, vol 17, pp 185–203, 1981.
  - cited ca. 15000 in 2020. This paper is often considered the first variational method in computer vision

Local and Global Approaches

Local optical flow approaches:

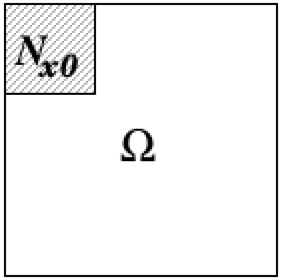
- the optical flow vector $u(x)$ at pixel $x = (x,y)^\top$ is computed by using only the image information in the local neighborhood of x.

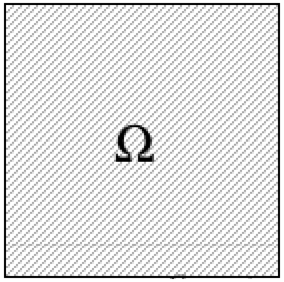
→ A famous instance is the local approach from **Lucas and Kanade**.

Global optical flow approaches:

- The optical flow field is computed simultaneously for the whole image domain Ω by solving a global optimization problem.
- These approaches are usually obtained by exploiting variational methods.

→ Probably best known is the **Horn and Schunck** approach.





Lucas-Kanade

Main idea:

The optical flow constraint $I_t + u^\top \nabla I = 0$ holds for all pixels in a (small) neighborhood $\mathcal{N}(x_0)$ with the same flow vector u around the current pixel x_0.

$$I_t(x_i, y_i, t) + I_x(x_i, y_i, t)\,\tilde{u} + I_y(x_i, y_i, t)\,\tilde{v} \approx 0 \quad \text{for} \quad x = (x_i, y_i)^\top \in \mathcal{N}(x_0)$$

Approach: (weighted least square fit)

Find the optical flow vector $u = (\tilde{u}, \tilde{v})^\top$ which minimizes the sum of the

Gaussian weighted (squared) deviations from the optical flow constraint.

$$\begin{aligned} J(u) &= \sum_{x \in \mathcal{N}(x_0)} G_\sigma(x - x_0) \underbrace{\left(I_t(x) + u^\top \nabla I(x)\right)^2}_{\text{constraint deviation}} \\ &= \sum_{x \in \mathcal{N}(x_0)} G_\sigma(x - x_0) \left(I_t^2(x) + 2u^\top \nabla I(x) I_t(x) + u^\top \nabla I(x) \nabla I(x)^\top u\right) \end{aligned}$$

Where G_σ denotes a Gaussian with variance sigma → Find an optimum $\dfrac{\partial J(u)}{\partial u} = 0$

Lucas-Kanade

Computing the derivative of J with respect to the flow u leads to the following condition for optimality:

$$\begin{aligned}\frac{\partial J(u)}{\partial u} &= \sum_{x\in\mathcal{N}(x_0)} \frac{\partial}{\partial u} G_\sigma(x-x_0)\left(I_t^2(x) + 2u^\top \nabla I(x) I_t(x) + u^\top \nabla I(x) \nabla I(x)^\top u\right) \\ &= \sum_{x\in\mathcal{N}(x_0)} G_\sigma(x-x_0)(2\nabla I(x) I_t(x) + 2\nabla I(x)\nabla I(x)^\top u) = 0\end{aligned}$$

The last equation allows to compute an u that belongs to an optimum of $J(u)$. Writing this component wise one has to solve the following 2×2-system:

$$\begin{pmatrix} \sum\limits_{x\in\mathcal{N}} G_\sigma(\tilde{x})(I_x)^2 & \sum\limits_{x\in\mathcal{N}} G_\sigma(\tilde{x}) I_x I_y \\ \sum\limits_{x\in\mathcal{N}} G_\sigma(\tilde{x}) I_y I_x & \sum\limits_{x\in\mathcal{N}} G_\sigma(\tilde{x})(I_y)^2 \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} -\sum\limits_{x\in\mathcal{N}} G_\sigma(\tilde{x}) I_t I_x \\ -\sum\limits_{x\in\mathcal{N}} G_\sigma(\tilde{x}) I_t I_y \end{pmatrix}.$$

$$G_\rho \star \begin{pmatrix} (\partial_{x_1} I)^2 & \partial_{x_1} I \partial_{x_2} I \\ \partial_{x_2} I \partial_{x_1} I & (\partial_{x_2} I)^2 \end{pmatrix} u = -G_\rho \star \begin{pmatrix} \partial_t I \partial_{x_1} I \\ \partial_t I \partial_{x_2} I \end{pmatrix}$$

(with $\tilde{x} = x - x_0$, $I_x = I_x(x) = I_x(x_i, y_j, t)$ and similar to I_y I_t)

Note: There is no unique solution in image regions with homogeneous image information or at image edges.

Inverse 2x2 Matrix:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{\mathrm{Det}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

with $\mathrm{Det} = a_{11}a_{22} - a_{21}a_{12}$

Notes on the Lucas and Kanade Approach:

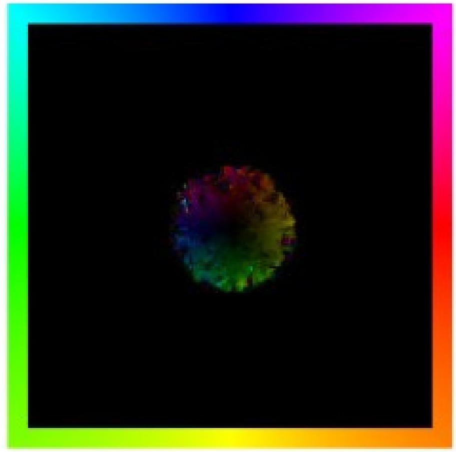
- **robust against noise** as it represents a least square fit with Gaussian weighted contributions
- **not invertible in image regions where the image gradient vanishes**
- in regions with one eigenvalue zero ($\mathrm{Det} = 0$) the aperture problem remains and no computation of the full optical flow is possible.

→ **robust, but no dense flow field**

Example – Lucas-Kanade

First image

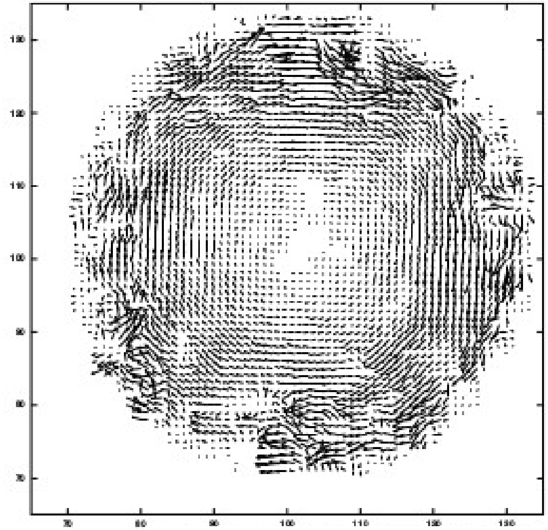
optical flow (color rep.)



second image

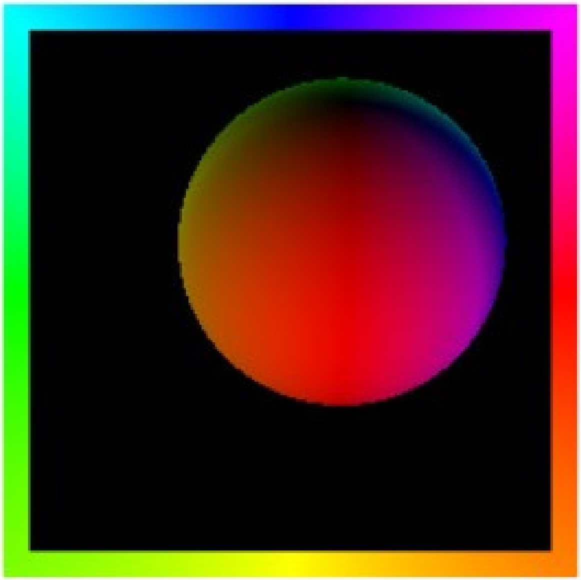
Original color image

flow vector field (zoom)



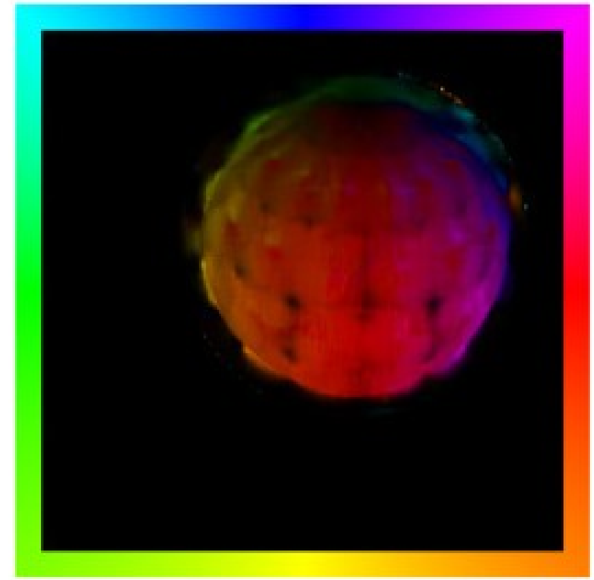
Example Lucas-Kanade

Ground truth

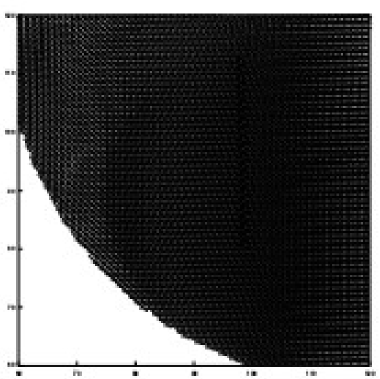


3x3 neighborhood

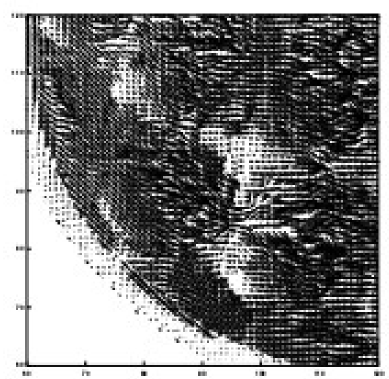
6x6 neighborhood



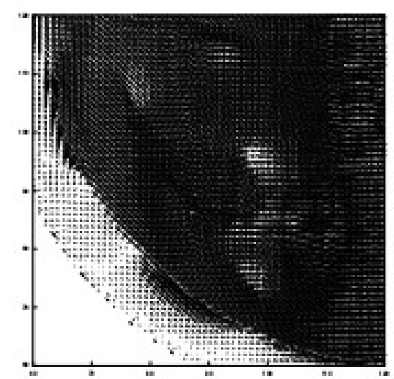
Ground truth



3x3 neighborhood



6x6 neighborhood



Variational Optical Flow Approach

Idea: Define an energy functional $E(u)$ which is minimal for the “best“ optical flow field u^*.

The solution is obtained within a variational framework. Typically $E(u)$ consists of a data term and a regularization term:

$$E(u) = E_{data}(u) + E_{regu}(u)$$

These terms can be seen as error measures that measure the deviation from certain constraints.

$$u^* = \underset{u}{\operatorname{argmin}}\, E(u) = \underset{u}{\operatorname{argmin}}(E_{data}(u) + E_{regu}(u))$$

Horn and Schunck incorporated the following 2 constraints into their optical flow approach

- **Data term**: The error in optical flow constraint equation should be small
- **Regularization term:** The optical flow field should be smooth

Horn and Schunck

Data term: Represents the sum/integral of all (squared) deviations from the optical flow constraint of the flow field u:

$$E_{data}(u) = \int_{\Omega} \underbrace{\left(I_t(x) + u^\top \nabla I(x)\right)^2}_{\text{constraint deviation}} dx$$

→ this term provides the link between the given **image data** and the desired velocity field

Regularisation term: Optical flow field should be smooth (no large changes).

$$E_{regu}(u) = \int_{\Omega} \alpha^2 \underbrace{\left[(\nabla u)^2 + (\nabla v)^2\right]}_{\text{smoothness}} dx$$

→ provides an additional constraint to make the problem unique solvable.

Note: The regularization term can be used to incorporate other constraints on the optical flow field too. (α^2 = relative importance of the two terms)

Variational Method

Variational Problem:

$$E(u) = \int_\Omega \underbrace{L(u, \frac{\partial u_i}{\partial x_i}, x)}_{\text{Lagrange function}} \; dx \rightarrow \text{minimum}$$

Lagrange Function:
(Horn and Schunck)
$u = (\tilde{u}, \tilde{v})^\top$

$$\begin{aligned} L &= \underbrace{\left(I_t(x) + u^\top \nabla I(x)\right)^2}_{\text{constraint deviation}} + \alpha^2 \underbrace{\left[(\nabla\tilde{u})^2 + (\nabla\tilde{v})^2\right]}_{\text{smoothness}} \\ &= (I_t + I_x\tilde{u} + I_y\tilde{v})^2 + \alpha^2 \left[\tilde{u}_x^2 + \tilde{u}_y^2 + \tilde{v}_x^2 + \tilde{v}_y^2\right] \end{aligned}$$

Euler-Lagrange Equations:
(optimality condition)

$$\frac{\partial L}{\partial \tilde{u}} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \tilde{u}_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial \tilde{u}_y} = 0$$

$$\frac{\partial L}{\partial \tilde{v}} - \frac{\partial}{\partial x}\frac{\partial L}{\partial \tilde{v}_x} - \frac{\partial}{\partial y}\frac{\partial L}{\partial \tilde{v}_y} = 0$$

Euler-Lagrange (Horn Schunck):
(vector equation)

$$\left(I_t(x) + u^\top \nabla I(x)\right) \nabla I - \alpha^2 \Delta u = 0$$

Variational Method

Euler-Lagrange (Horn Schunck):

$$\begin{aligned} I_t I_x + I_x^2 \tilde{u} + I_x I_y \tilde{v} - \alpha^2 \Delta \tilde{u} &= 0 \\ I_t I_y + I_x I_y \tilde{u} + I_y^2 \tilde{v} - \alpha^2 \Delta \tilde{v} &= 0 \end{aligned}$$

Approximating the Laplacian:

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

$$\Delta \tilde{u} \approx \bar{u} - \tilde{u} \quad \text{and} \quad \Delta \tilde{v} \approx \bar{v} - \tilde{v}$$

Approximation Δu:

$$\frac{1}{12}\underbrace{\begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 0 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}}_{\bar{u}} - \underbrace{\begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array}}_{u}$$

Inserting these approximations one finds the following 2 × 2 system

$$\begin{pmatrix} (I_x^2 + \alpha^2) & I_x I_y \\ I_x I_y & (I_y^2 + \alpha^2) \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} \alpha^2 \bar{u} - I_t I_x \\ \alpha^2 \bar{v} - I_t I_y \end{pmatrix}.$$

which can be easily solved for $\tilde{u}$ and $\tilde{v}$:

$$\begin{aligned} (\alpha^2 + I_x^2 + I_y^2)\tilde{u} &= (I_y^2 + \alpha^2)\bar{u} - I_x I_y \bar{v} - I_t I_x \\ (\alpha^2 + I_x^2 + I_y^2)\tilde{v} &= -I_x I_y \bar{u} + (I_x^2 + \alpha^2)\bar{v} - I_t I_x \end{aligned}$$

Variational Method

Subtracting the terms $(\alpha^2 + I_x^2 + I_y^2)\bar{u}$ from the first and $(\alpha^2 + I_x^2 + I_y^2)\bar{v}$ from the second equation one gets:

$$\begin{aligned}\tilde{u} &= \bar{u} - I_x[I_x\bar{u} + I_y\bar{v} - I_t]/(\alpha^2 + I_x^2 + I_y^2) \\ \tilde{v} &= \bar{v} - I_y[I_x\bar{u} + I_y\bar{v} - I_t]/(\alpha^2 + I_x^2 + I_y^2)\end{aligned}$$

This gives rise to the following simple iterative gradient descent method (Coupled Gauss-Seidel relaxation):

$$\begin{aligned}\tilde{u}^{n+1} &= \bar{u}^n - I_x[I_x\bar{u}^n + I_y\bar{v}^n - I_t]/(\alpha^2 + I_x^2 + I_y^2) \\ \tilde{v}^{n+1} &= \bar{v}^n - I_y[I_x\bar{u}^n + I_y\bar{v}^n - I_t]/(\alpha^2 + I_x^2 + I_y^2)\end{aligned}$$

(The superscript $n + 1$ indicates the next iteration.)

Discretization with Finite Elements

- Energy functional is strictly convex
  → exploit FEM method

- Regular triangulation

- Attach to each pixel position a piecewise linear basis function Φ

$$u_1(x_1, x_2) = \sum_{i=1}^{N} u_i \phi_i(x_1, x_2) \ ,$$

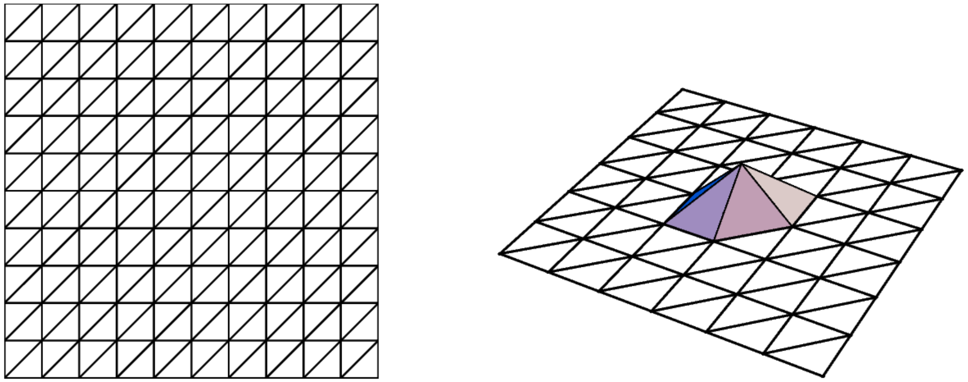


Figure 2.5: Left: Uniform triangulation of the image domain Ω. **Right:** Basis function $\phi_i(x_1, x_2)$ that belongs to a pixel position i.

→ sparse, positive definite linear system

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot A \begin{pmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{pmatrix} = b \cdot \begin{pmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{pmatrix} \ , \quad \forall \tilde{u}_1, \tilde{u}_2$$

$$A \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = b$$

- solve with corresponding iterative solver parallelization by domain decomposition

Discretization and Solution

The 2Nx2N A matrix factorizes into

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^{\top} & A_{22} \end{pmatrix},$$

with

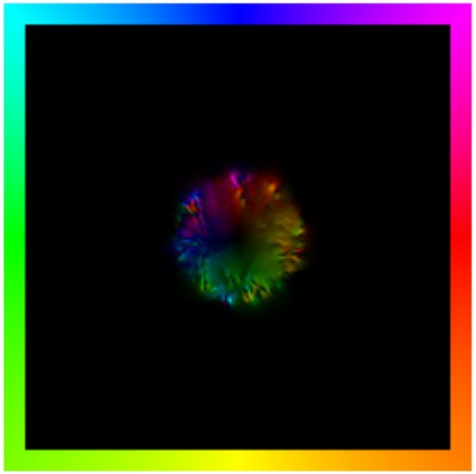
$$\begin{aligned}
(A_{11})_{k,l} &= a\big((\phi_k, 0)^{\top}, (\phi_l, 0)^{\top}\big) \\
(A_{12})_{k,l} &= a\big((\phi_k, 0)^{\top}, (0, \phi_l)^{\top}\big) \\
(A_{22})_{k,l} &= a\big((0, \phi_k)^{\top}, (0, \phi_l)^{\top}\big) \ .
\end{aligned}$$

and

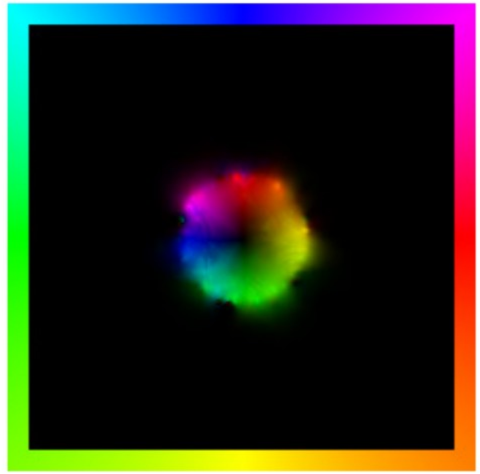
$$\begin{aligned}
(b_1)_k &= b\big((\phi_k, 0)^{\top}\big) \\
(b_2)_k &= b\big((0, \phi_k)^{\top}\big)
\end{aligned}$$

Horn and Schunck

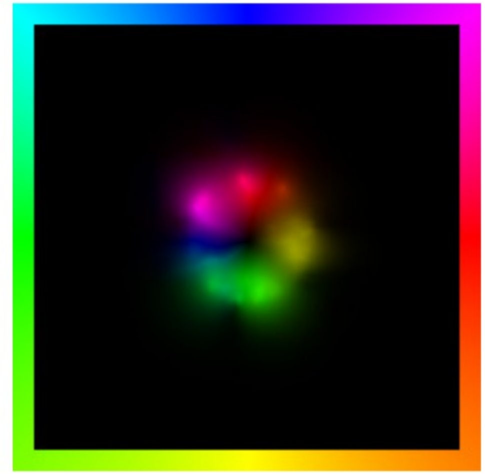
$\alpha = 0.1$



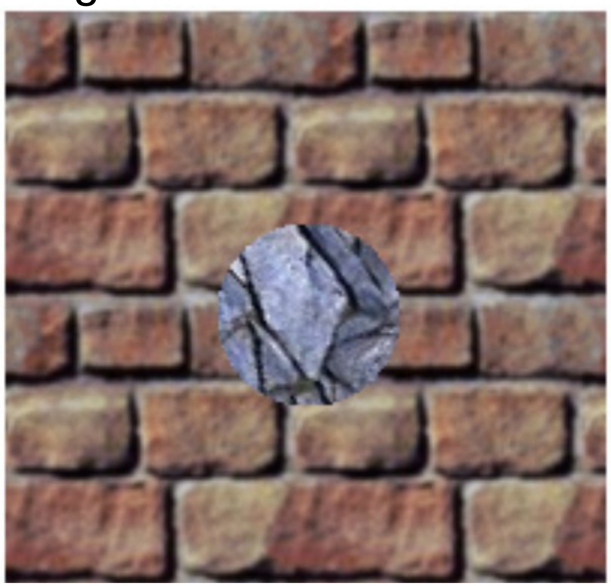
$\alpha = 0.5$



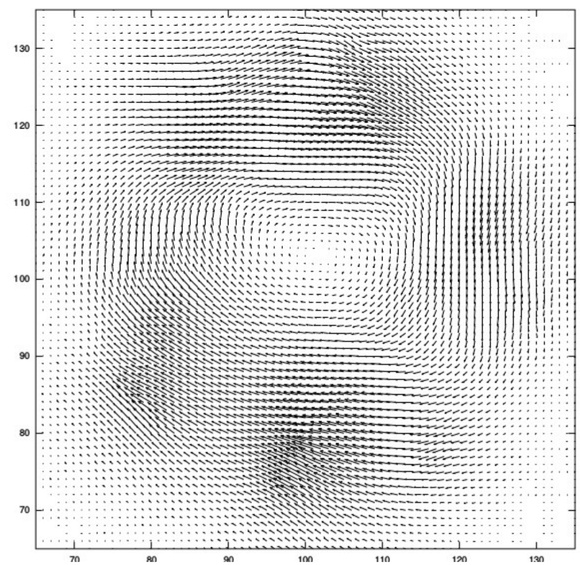
$\alpha = 0.8$



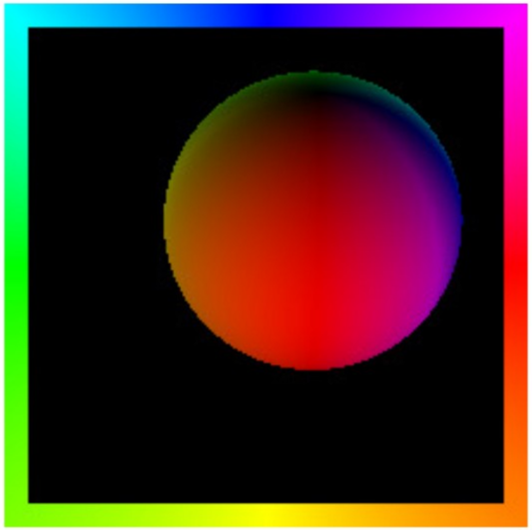
Original



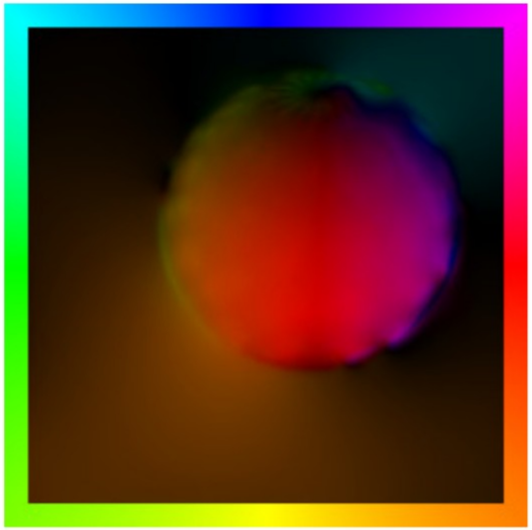
optical flow field (zoom in)



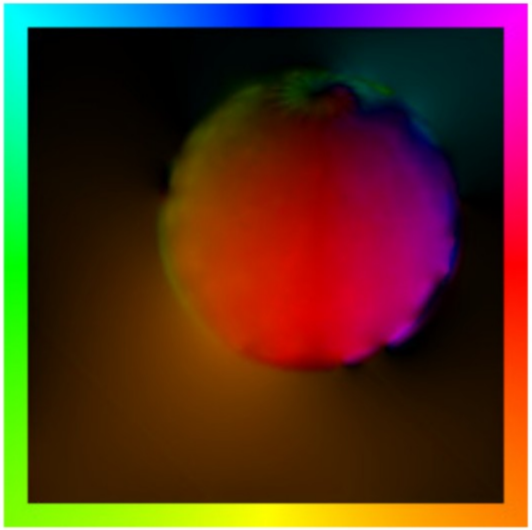
Horn and Schunck - Example

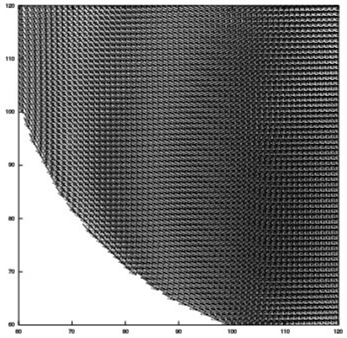


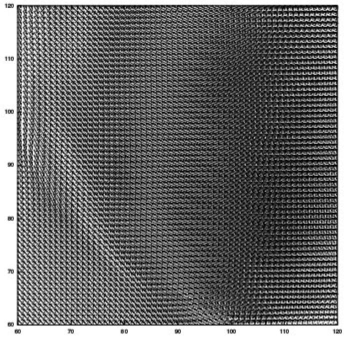
$\alpha = 0.5$

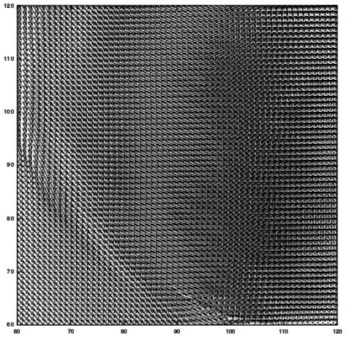


$\alpha = 0.8$









Horn and Schunk

Notes on the Horn and Schunck approach:

- the smoothness term **penalizes** non-smooth flow fields (spatial coherence assumption) and avoids an ill-posed optical flow formulation
- the smoothness term **helps to fill in** flow field information from the neighborhood to regions with no image gradient (fill-in effect)
- more sensitive to noise than local approaches
- the approach is **convex** and therefore converges to a unique solution.
- an increasing value of α forces the vector field to become smoother
- **Motion boundaries:** The homogeneous regularization at motion boundaries leads to a **smoothing over discontinuities** - thus sharp contours are lost.
- **Occlusions:** The smoothness assumption is clearly violated across occluded edges where the optical flow field changes abruptly.

→ **results in a dense flow field, more sensitive to noise**

Error Measures

For a quantitative comparison typically the following measures are used:

1) The **root mean square error (RMSE)**

$$RMSE(u_o, u_e) = \frac{1}{|\Omega|} \int\limits_{\Omega} \sqrt{(u_o - u_e)^2} dx$$

- $u_o = (u_{o_1}, u_{o_2}, 1)^\top$ the original (ground truth) optical flow vectors
- $u_e = (u_{e_1}, u_{e_2}, 1)^\top$ the estimated (computed) optic flow vectors
- (the time dimension is set to 1 corresponding to the distance of one frame)

2) The **average angular error (AAE)**

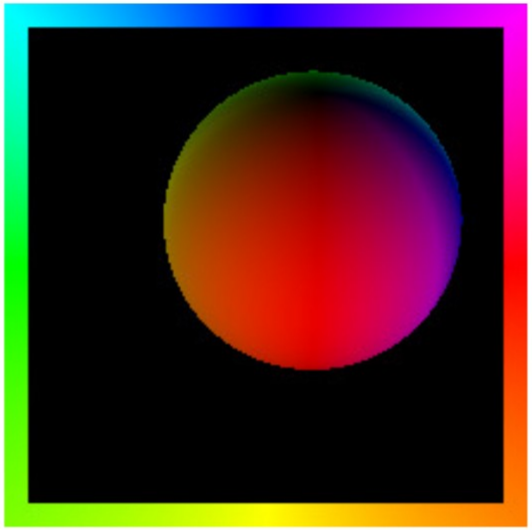
$$AAE(u_o, u_e) = \frac{1}{|\Omega|} \int\limits_{\Omega} \arccos\left(\frac{u_o \cdot u_e}{|u_o||u_e|}\right) dx,$$

This measure can also be used to provide accuracy measures for optical flow results.

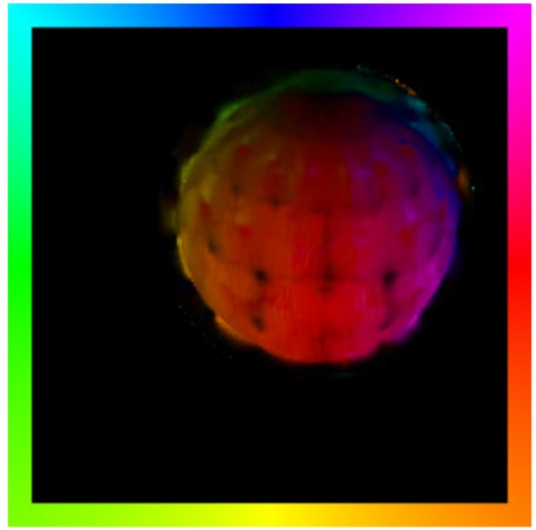
→ problematic areas are revealed when error images are plotted

Error Measure - RMS

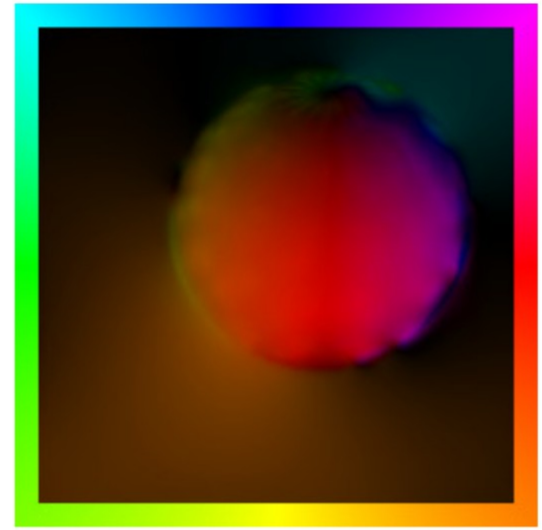
Ground truth



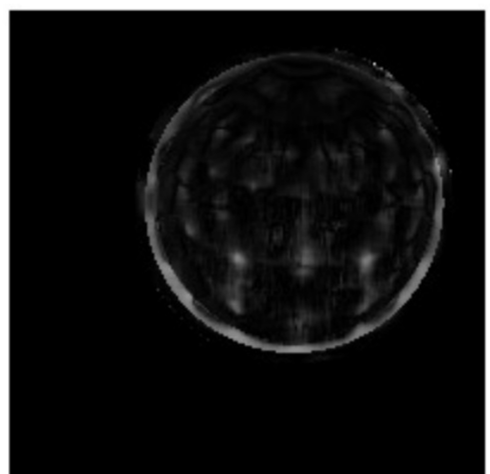
LK flow (6x6 mask)



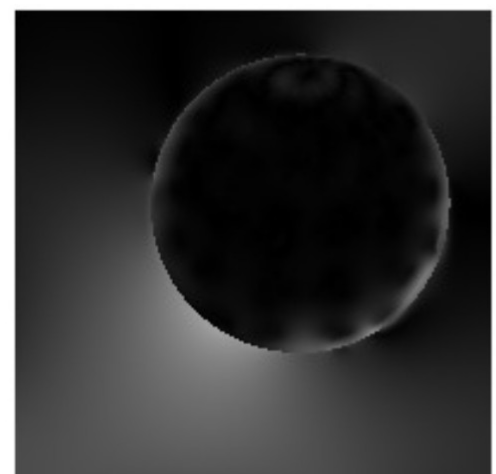
HS flow $(\alpha = 10.0)$



LK RMS-error



HS RMS-error



Large Displacements

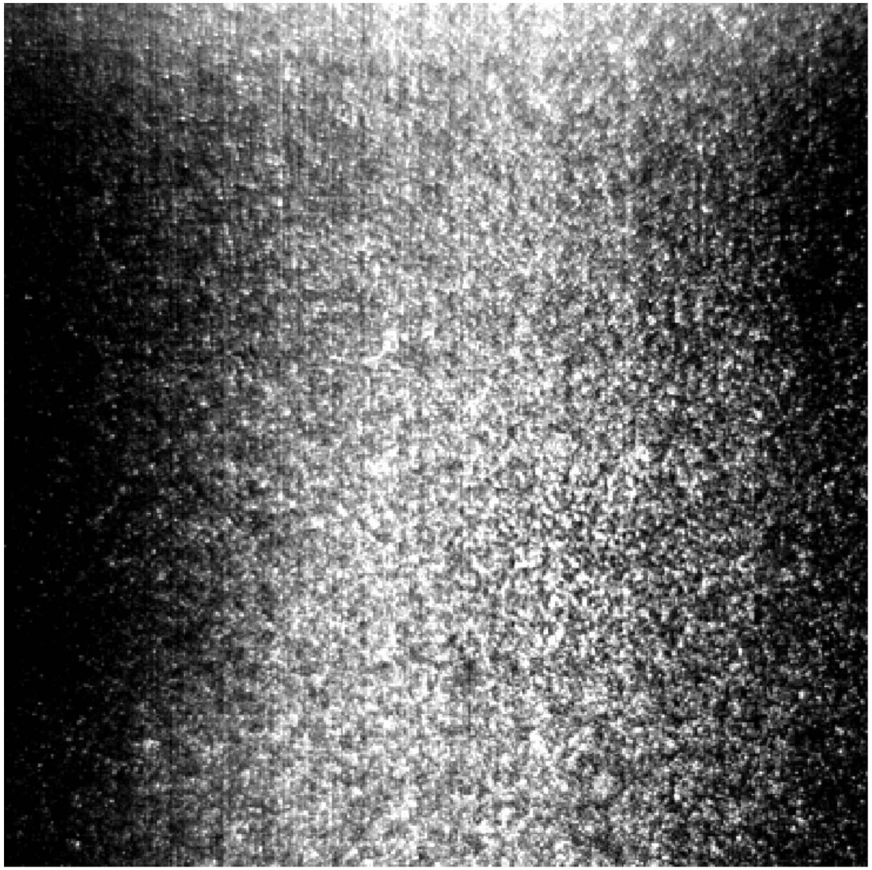
... to cope with large displacements

Standard optical flow approaches give good results only for offsets < 1-2 pixels. Why?

Temporal aliasing
Taylor series linearization:

$\rightarrow$ Use image pyramid for displacements

up to 20 px.



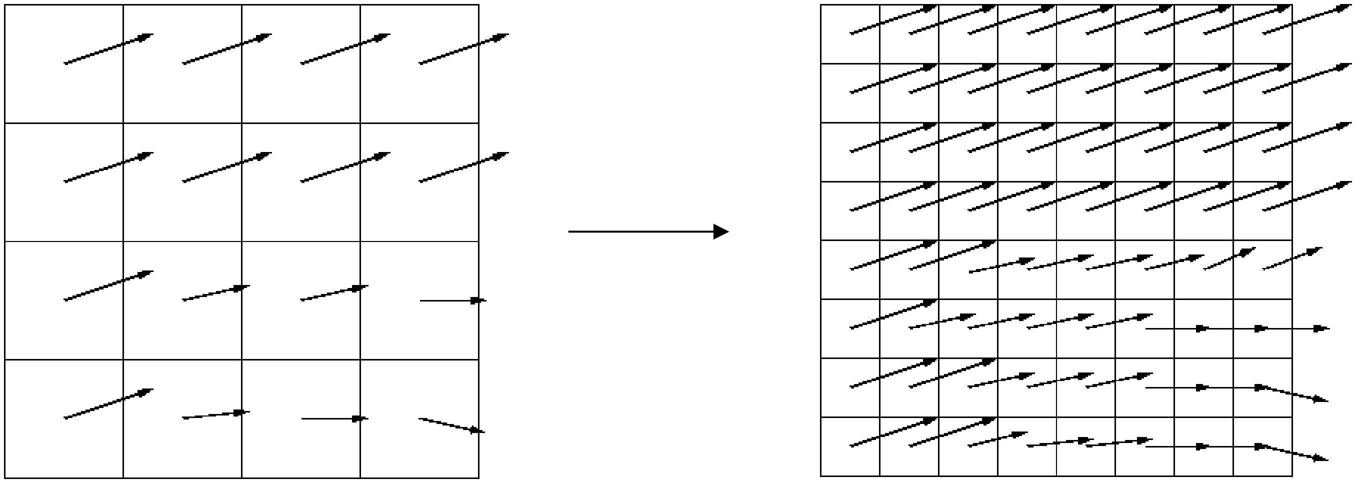
Build image pyramid

- Low-pass filtering
- Subsampling

Coarse-to-fine Strategy

From coarse to fine:

- Compute motion field (variational approach)
- Project to next resolution level



Warp image! stabilize positions of particles over time

Large Displacements

Techniques that help to cope with larger displacements:

- image pyramids (hierarchical scaling scheme)
- warping (warp images closer to each other using computed optical flow)

Example:
Brox, Bruhn, Papenberg, Weickert:
High accuracy optical flow estimation
based on a theory for warping,
ECCV 2004.

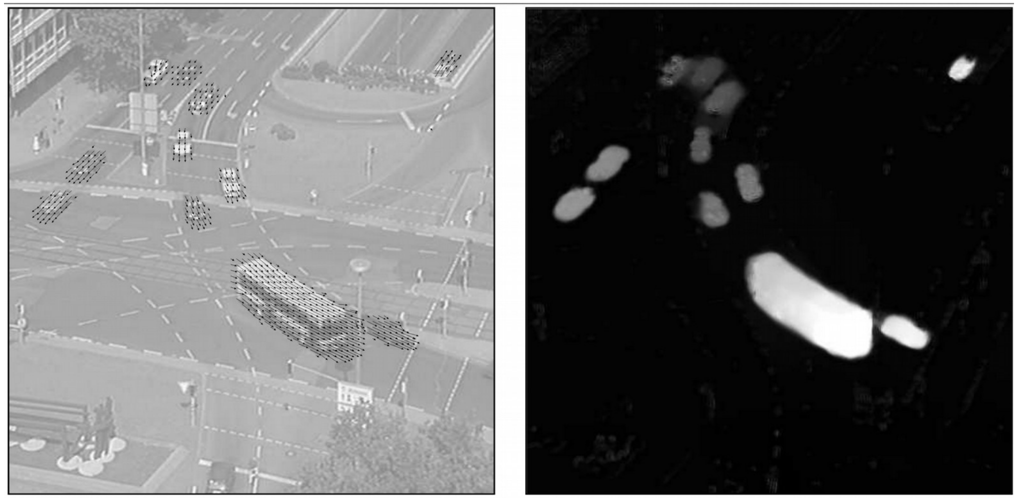
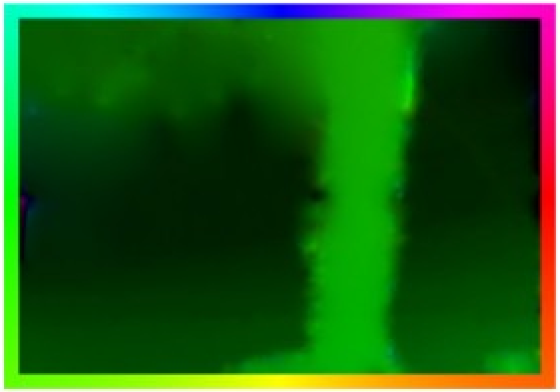


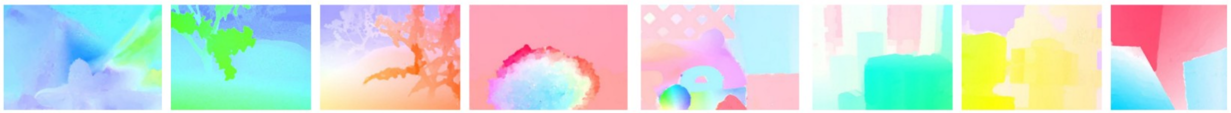
Image 1

image2

optical flow



Further Advances



Measure	$\rho_\mathbf{F}$	$g(\lvert\nabla I\rvert)$	Dimetrodon	Grove2	Grove3	Hydrangea	RubberWhale	Urban2	Urban3	Venus
EPE	–	–	0.19	0.15	0.67	0.15	0.09	0.32	0.63	0.26
EPE	×	–	0.28	0.15	0.65	0.15	0.09	**0.29**	0.49	0.26
EPE	–	×	**0.19**	0.15	0.58	0.15	**0.08**	0.32	0.60	0.26
EPE	×	×	**0.19**	**0.14**	**0.56**	0.15	**0.08**	**0.29**	**0.45**	**0.25**
rel-$\rho_\mathbf{F}$	×	×	0.11	0.01	0.02	0.21	0.29	0.01	0.01	0.01

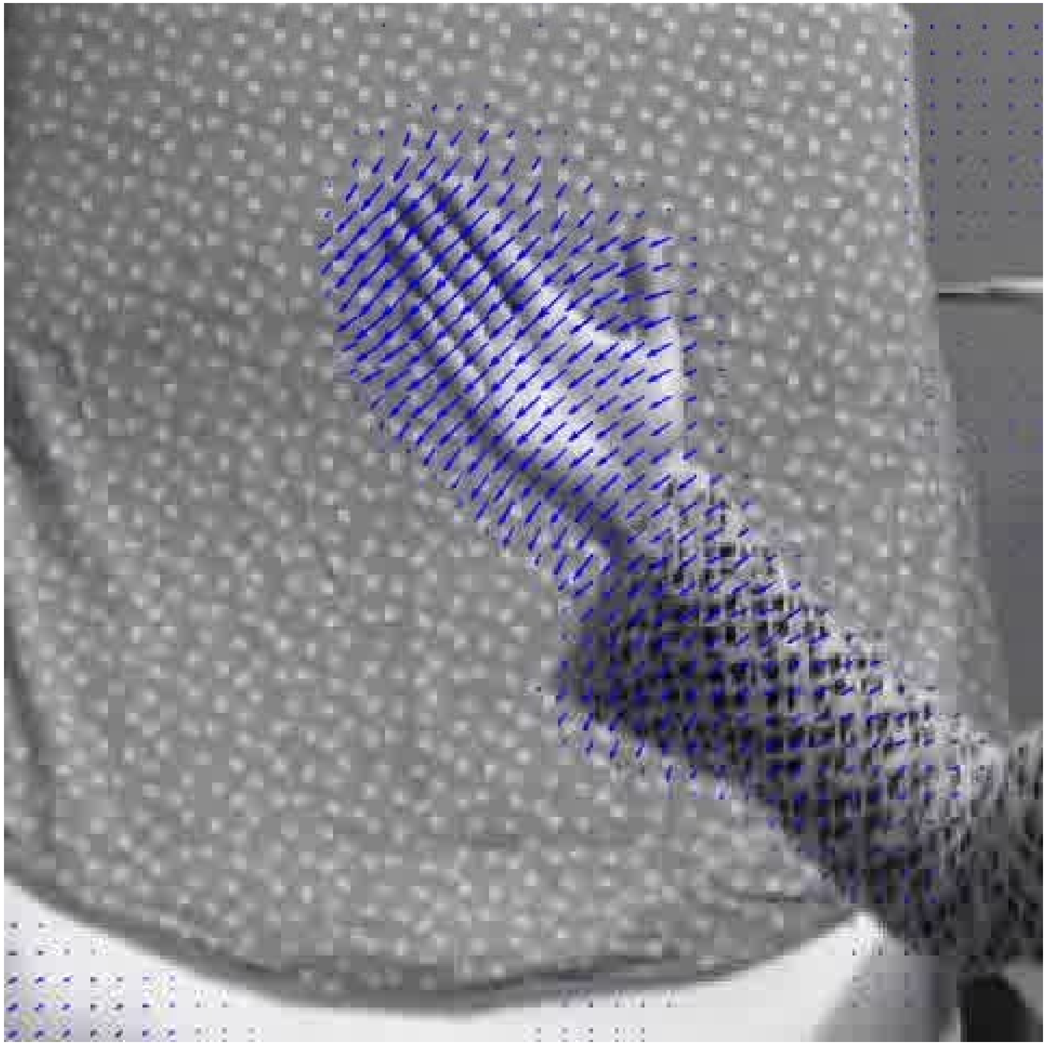
Table 1. Evaluation results on the Middlebury training data. The proposed regularizers, $\rho_\mathbf{F}$ and $g(|\nabla I|)$, systematically improve the optic flow estimates. In the table, EPE denotes the average end-point error of the obtained optic flow field and rel-$\rho_\mathbf{F} = \int_\Omega \rho_\mathbf{F}(v,x)/\|v\|\, d^2x$ is the average relative epipolar line distance. See Table 2 for execution times on the training data set.

Figure 1. Optical flow estimation in a dynamic scene. The lower images show the $\rho_\mathbf{F}$ constraint deviation and the edge image $|\nabla I|$. Note the accurate flow estimation for the approaching car and the detection of independent motion for the distant pedestrian.

Figure 2. Optical flow estimation for a scene with a running person. Even the large displacement of the right foot is correctly matched. The lower images show the $\rho_\mathbf{F}$ constraint deviations which clearly identify the person and the edge image $|\nabla I|$.

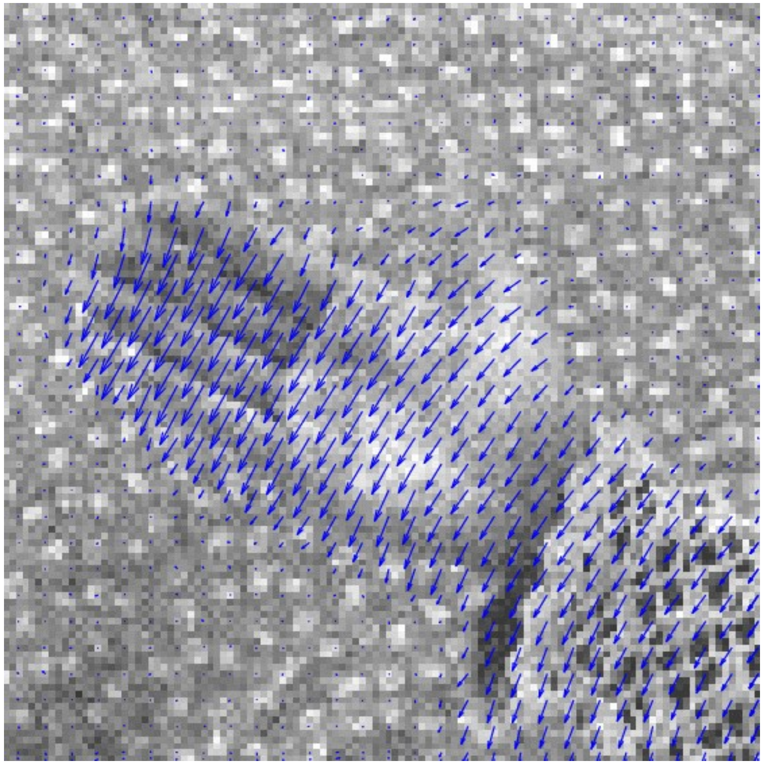
Wedel, Andreas & Cremers, Daniel & Pock, Thomas & Bischof, Horst. (2009). Structure- and Motion-adaptive Regularization for High Accuracy Optic Flow. Proceedings of the IEEE International Conference on Computer Vision. 1663 - 1668.

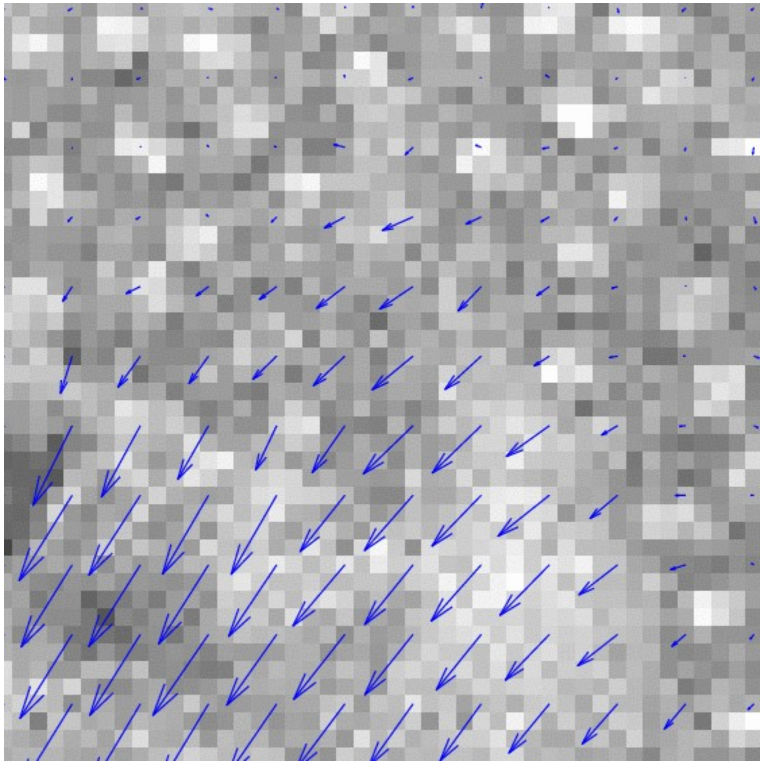
Motion Estimation and Tracking



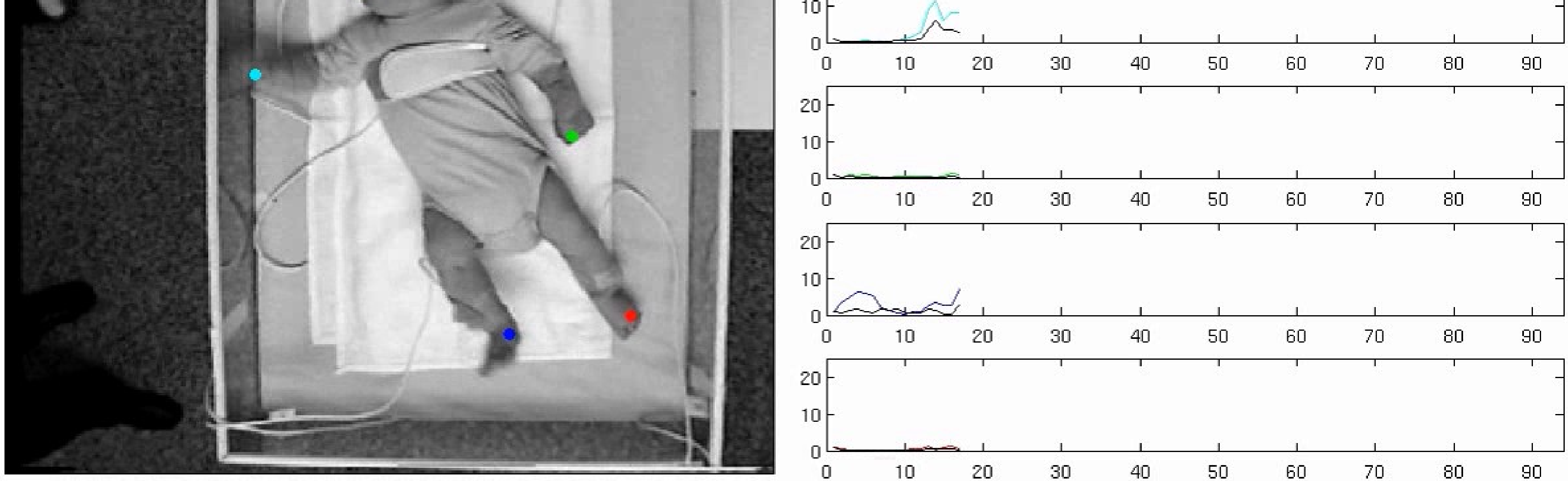
Motion Estimation - Noise

Control based Optical Flow Estimation
Results: Gaussian Noise sigma = 20



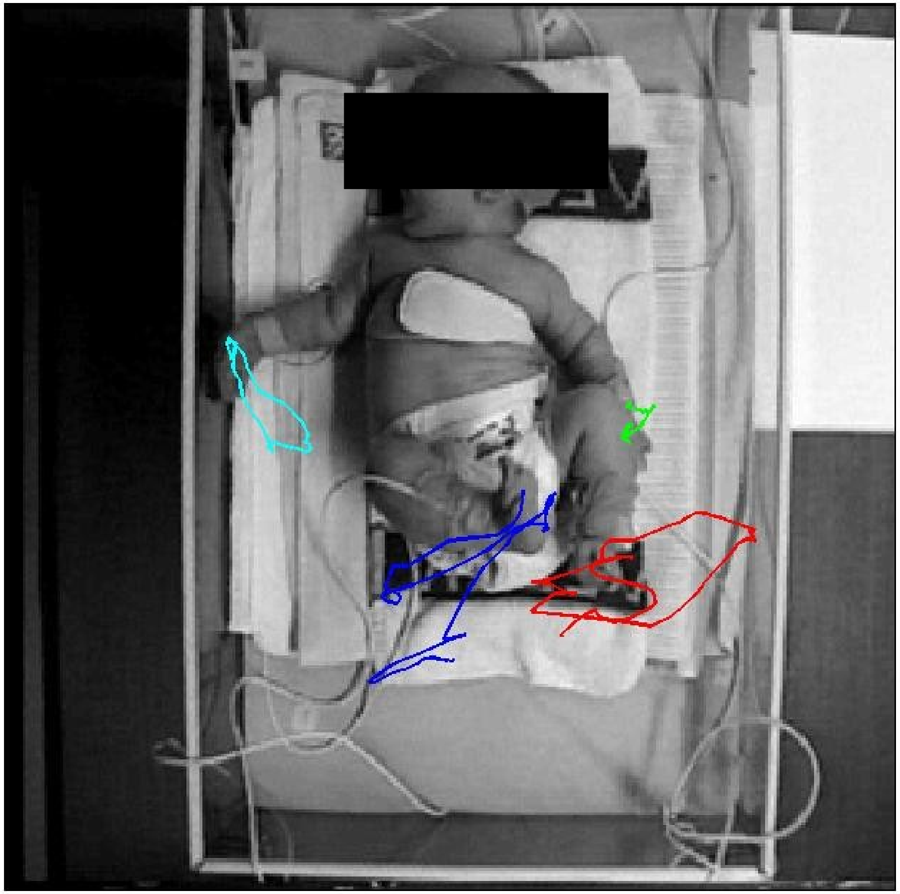


Stahl, Annette; Aamo, Ole Morten. (2011) A new Framework for Motion Estimation in Image Sequences using Optimal Flow Control. Informatics in Control, Automation and Robotics Revised and Selected Papers from the International Conference on Informatics in Control, Automation and Robotics 2010.



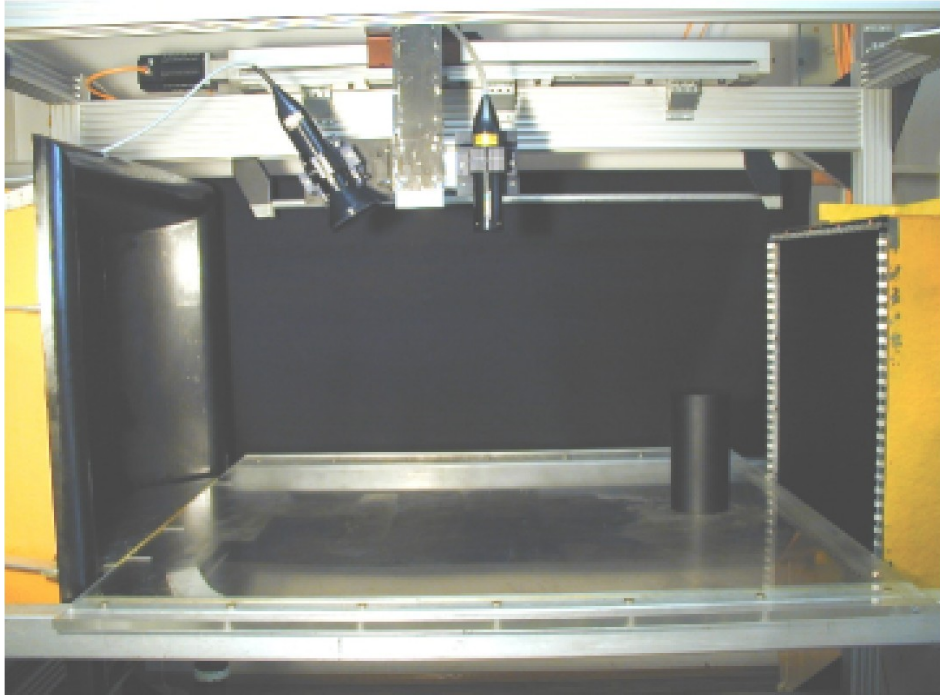
Stahl, Annette; Schellewald, Christian; Stavdahl, Øyvind; Aamo, Ole Morten; Adde, Lars; Kirkerød, Harald. (2012)
An Optical Flow-Based Method to Predict Infantile Cerebral Palsy. IEEE transactions on neural systems and rehabilitation engineering. vol. 20 (4).

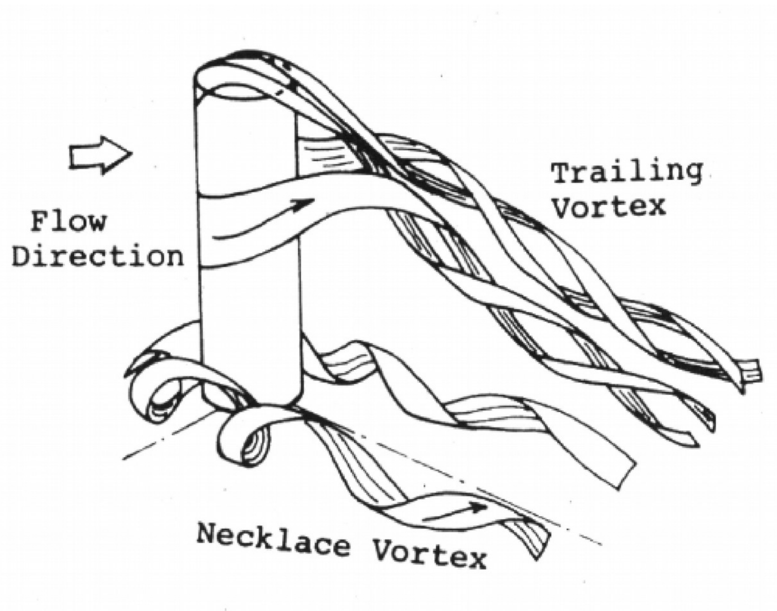
Motion Estimation and Tracking



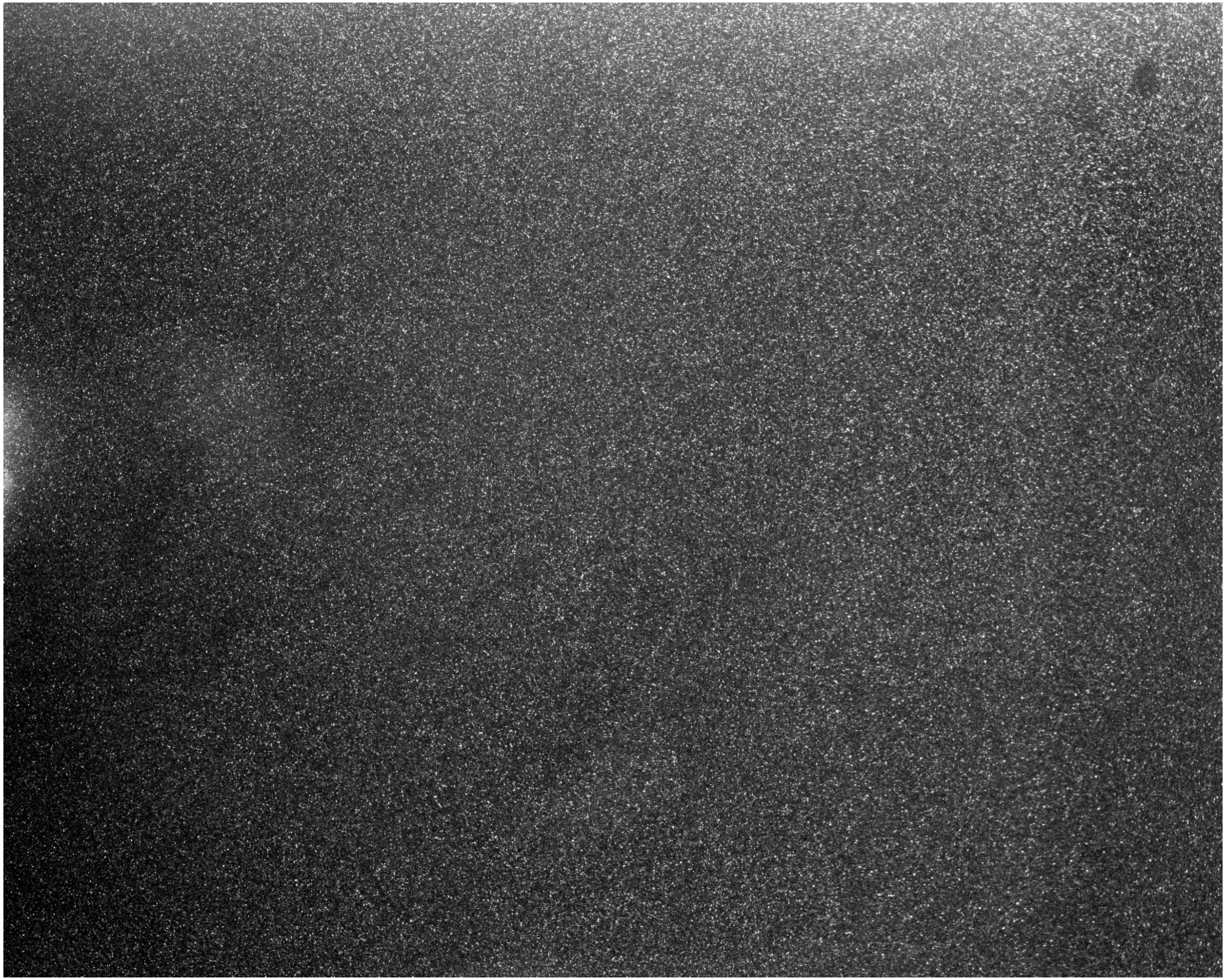
Stahl, Annette; Schellewald, Christian; Stavdahl, Øyvind; Aamo, Ole Morten; Adde, Lars; Kirkerød, Harald. (2012)
An Optical Flow-Based Method to Predict Infantile Cerebral Palsy. IEEE transactions on neural systems and rehabilitation engineering. vol. 20 (4).

Particle Image Velocimetry

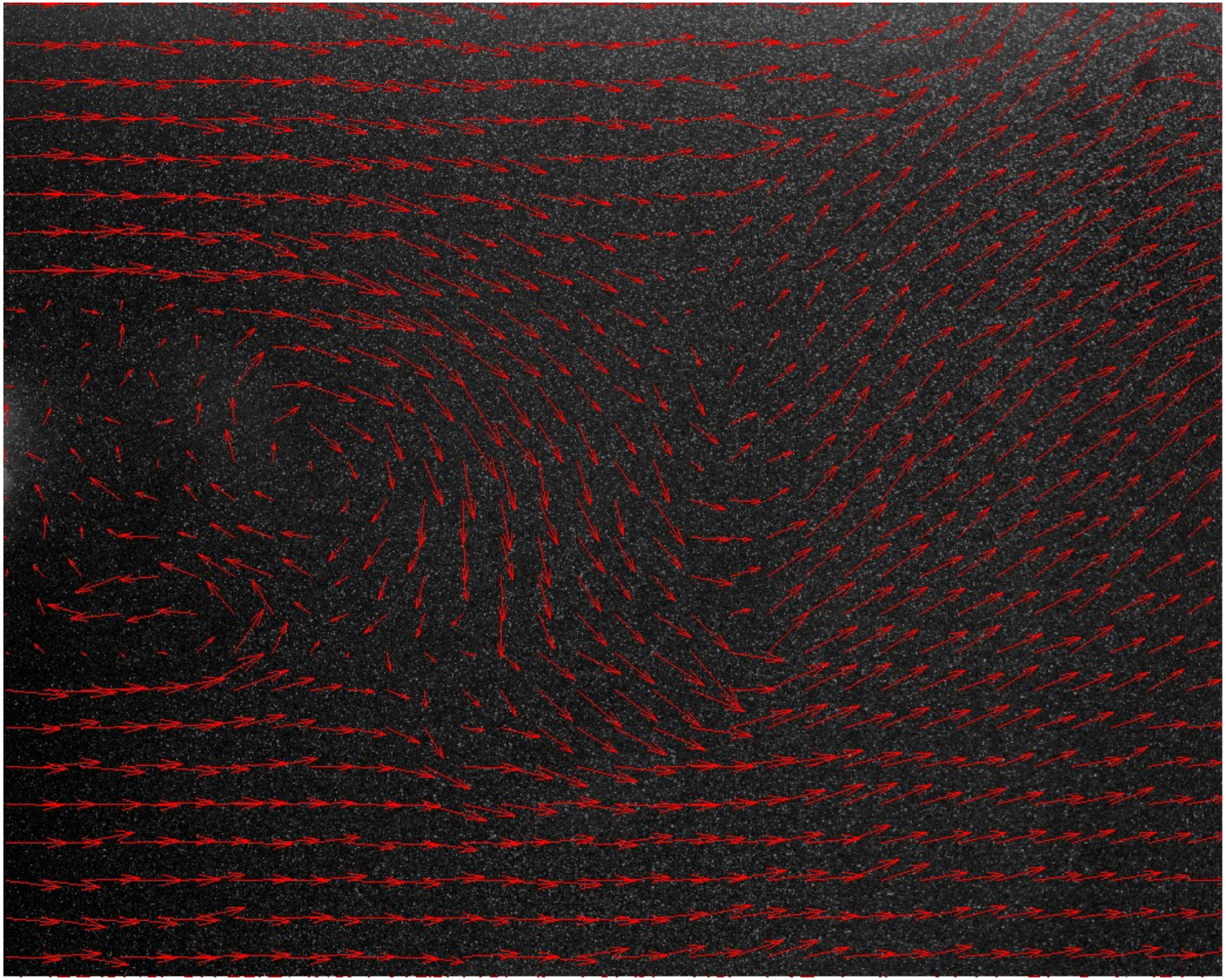




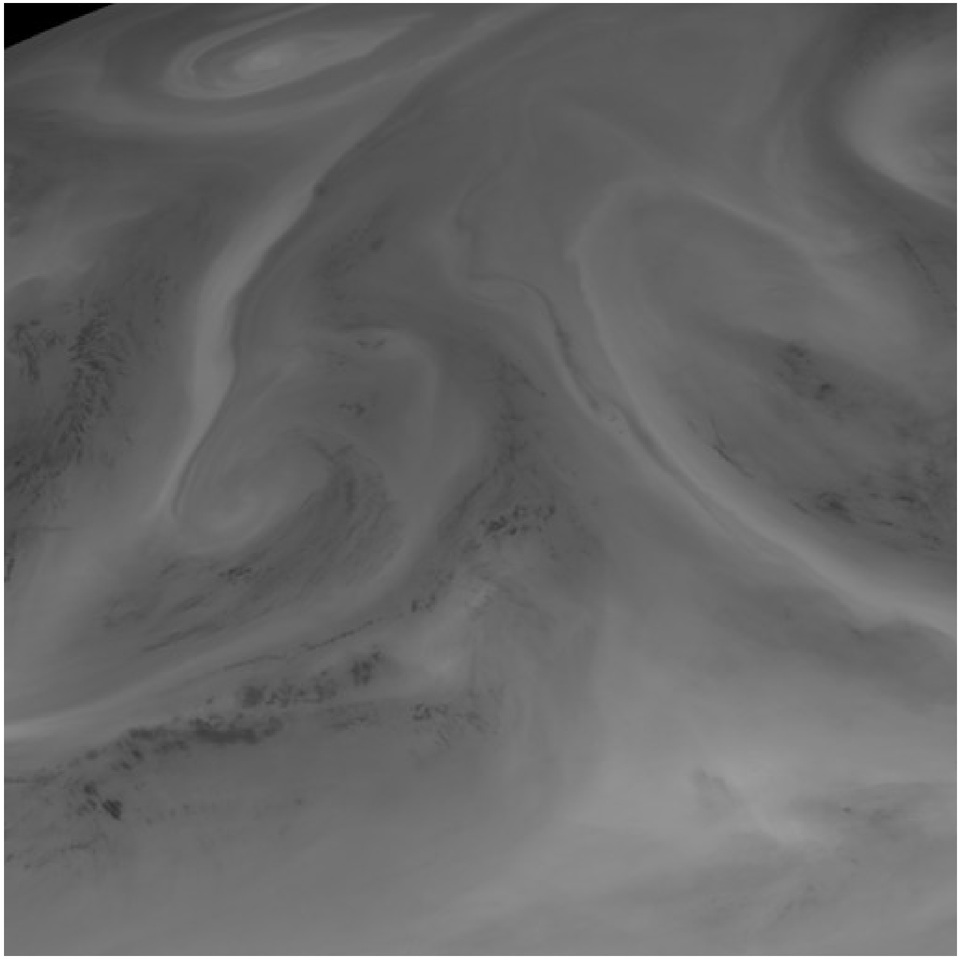
Particle Image Velocimetry

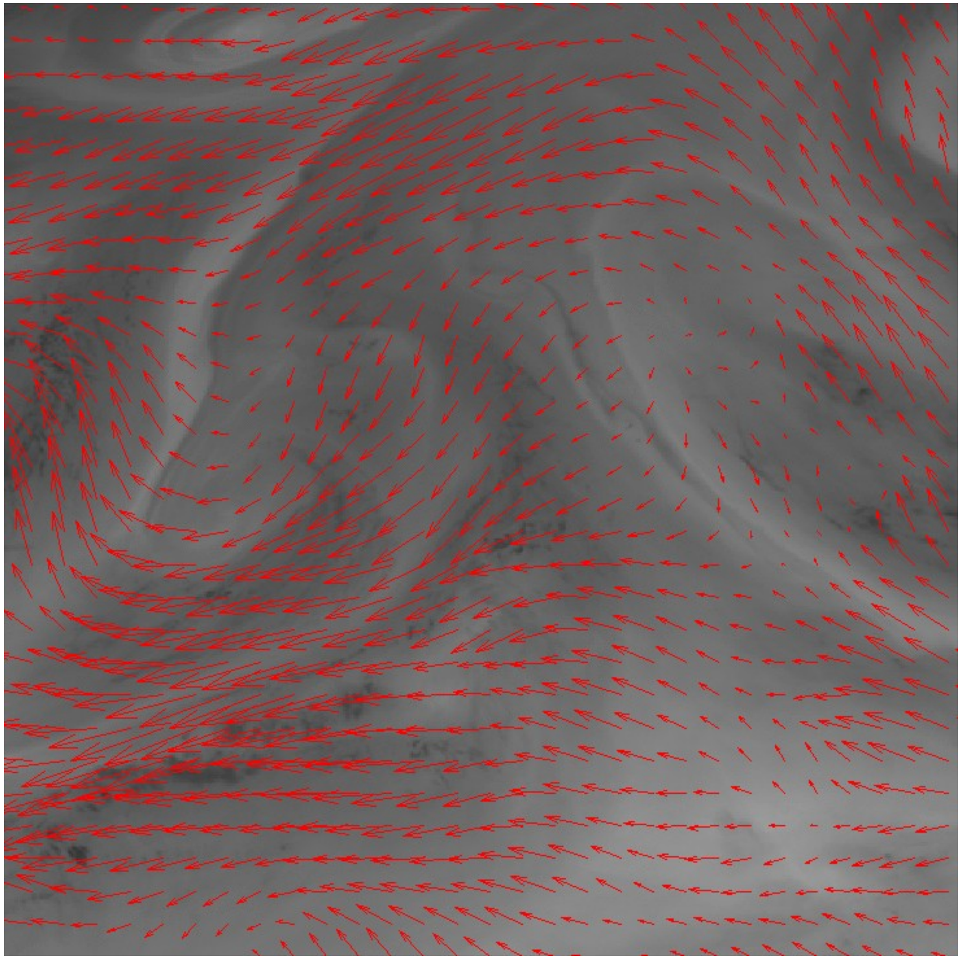


PIV using Variational Optical Flow Estimation

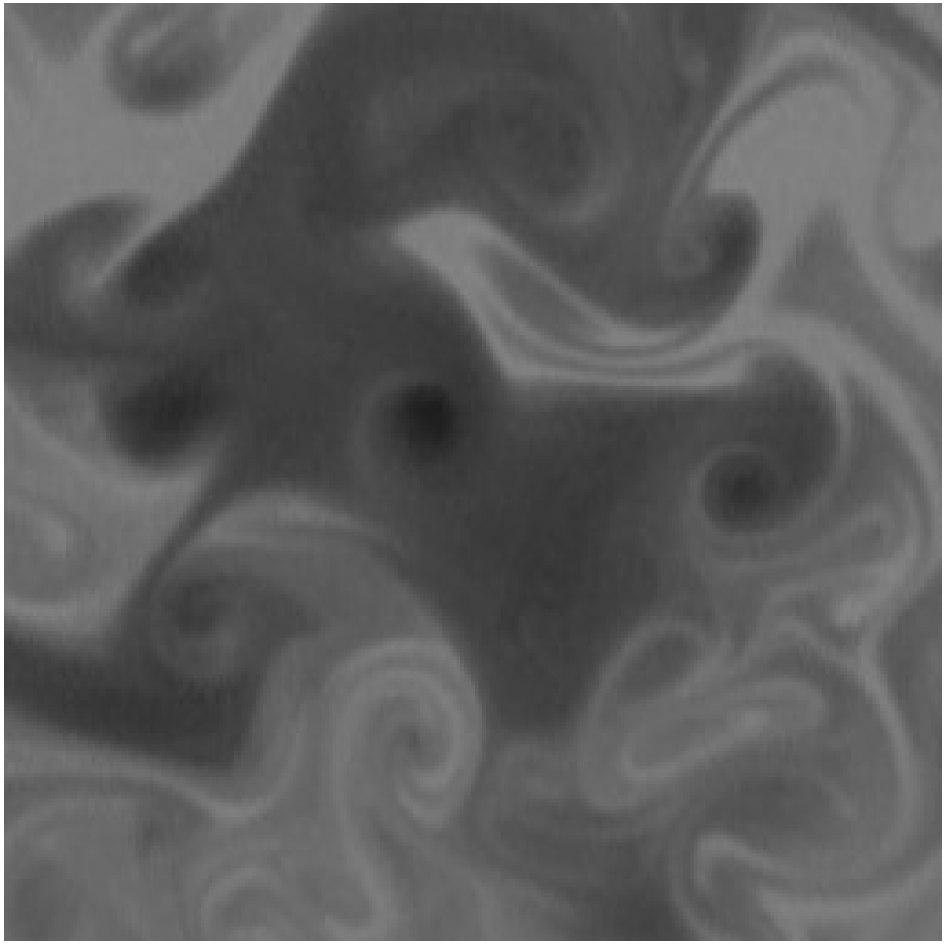


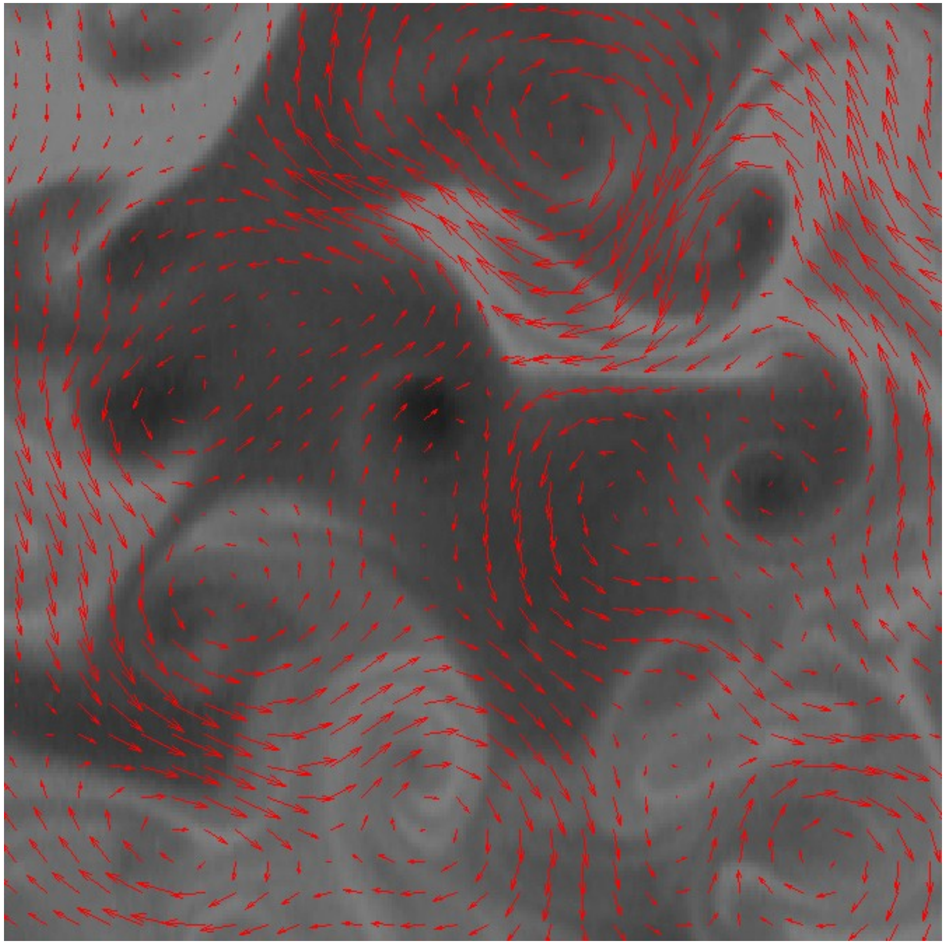
Variational Flow Control - Navier-Lamé



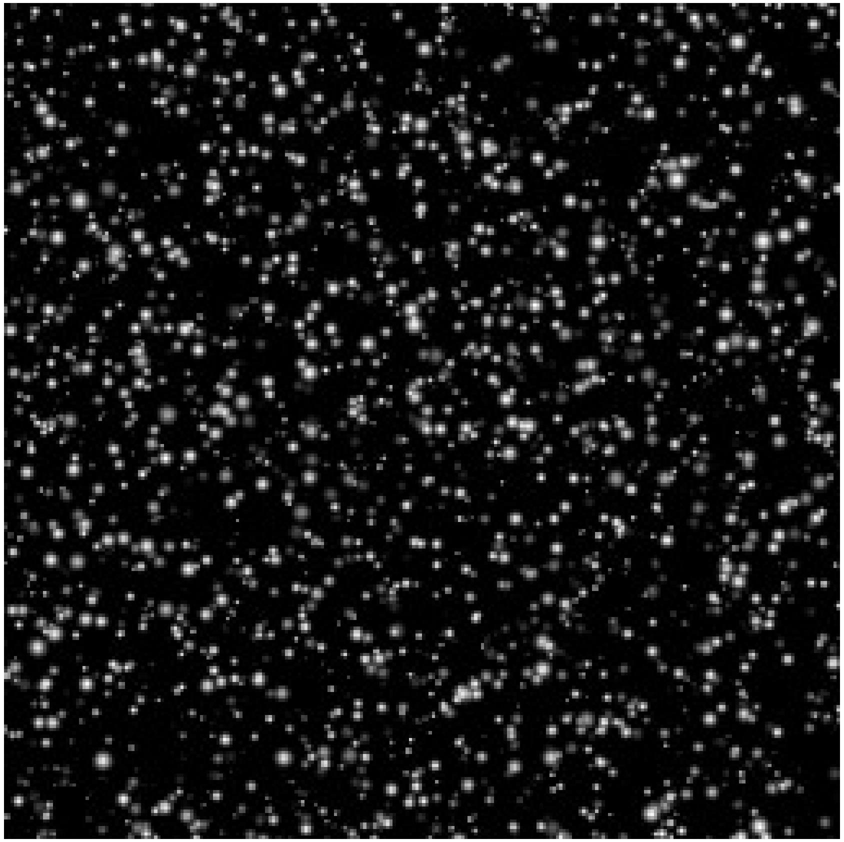


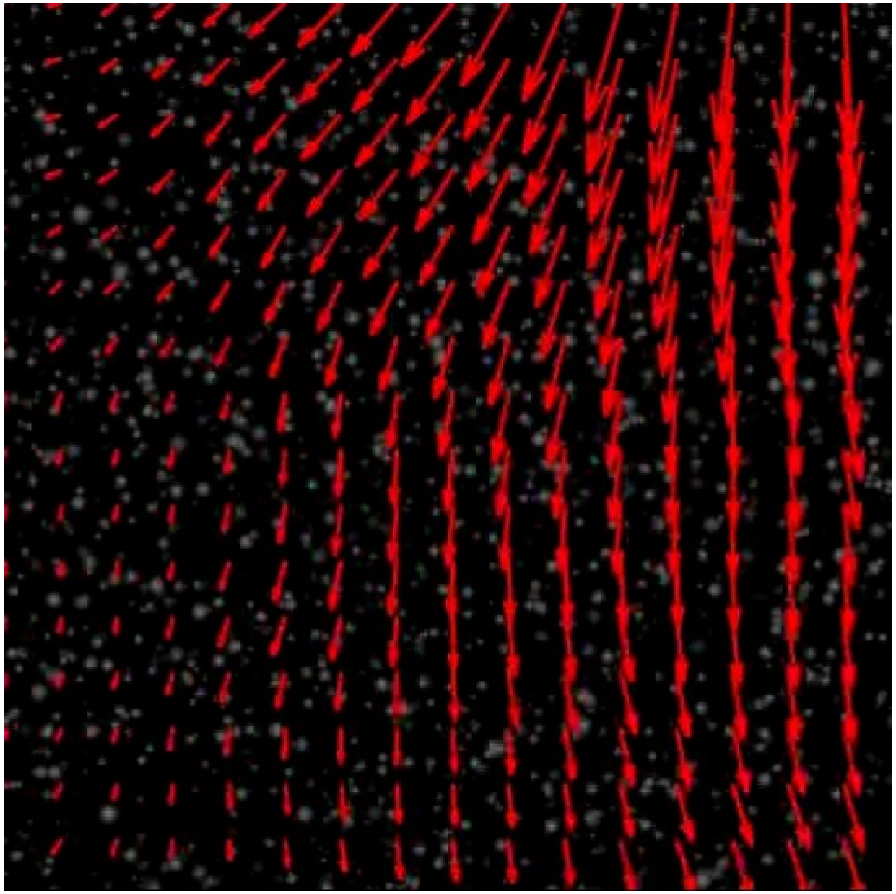
Variational Flow Control – Stokes Flow





PIV - Spatio-Temporal Extension



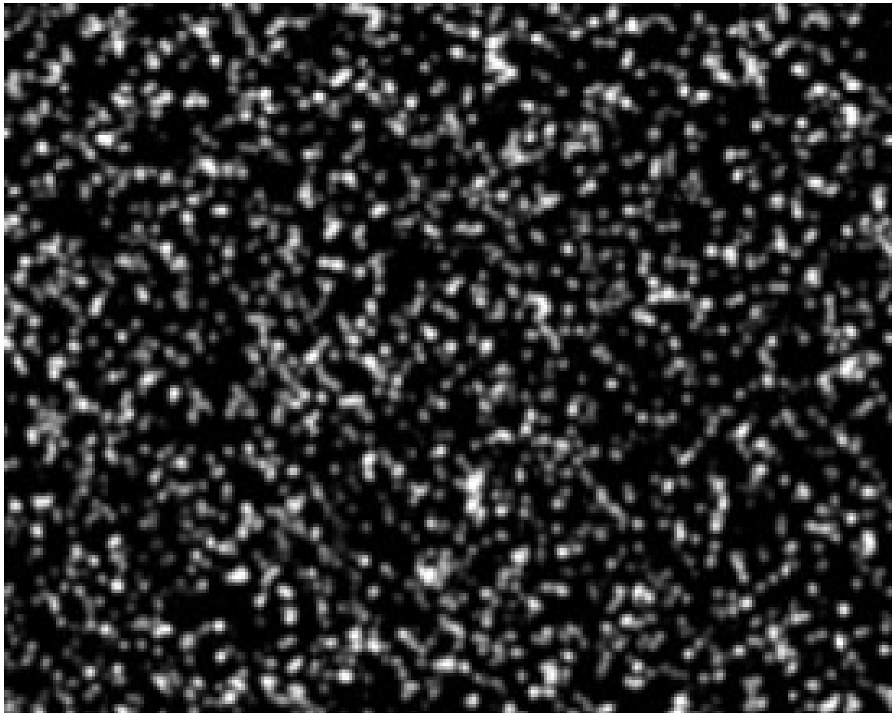


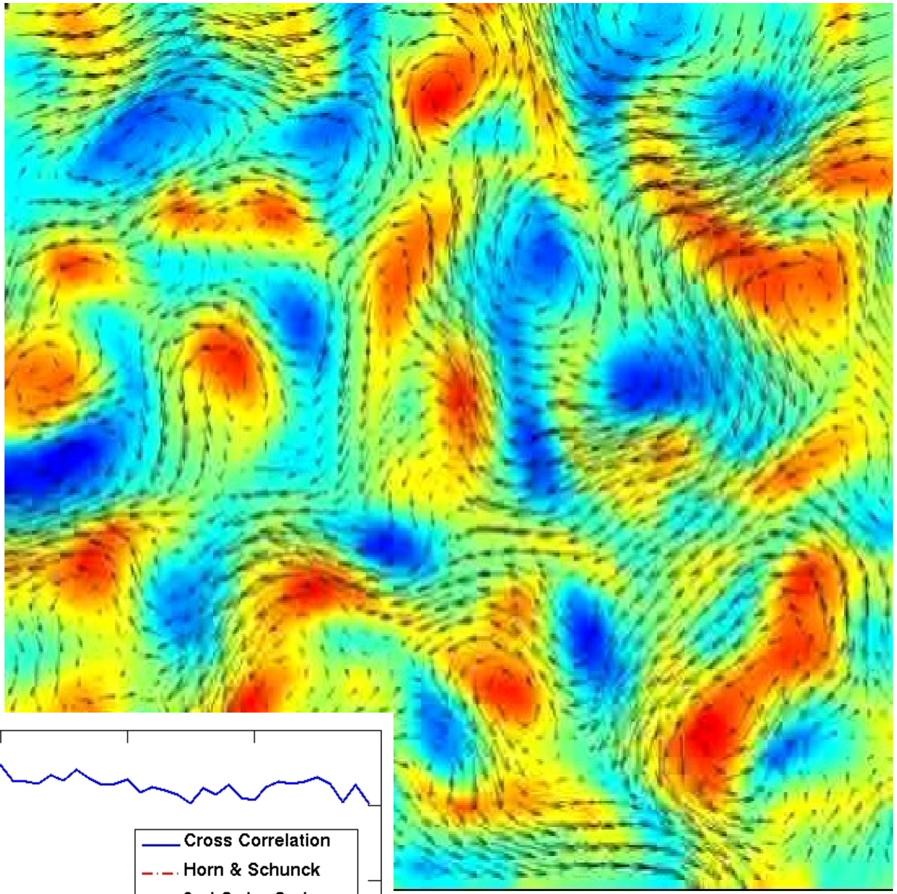
Weickert and Schnoerr proposed a modification of the 2D Horn and Schunck approach to three dimensions

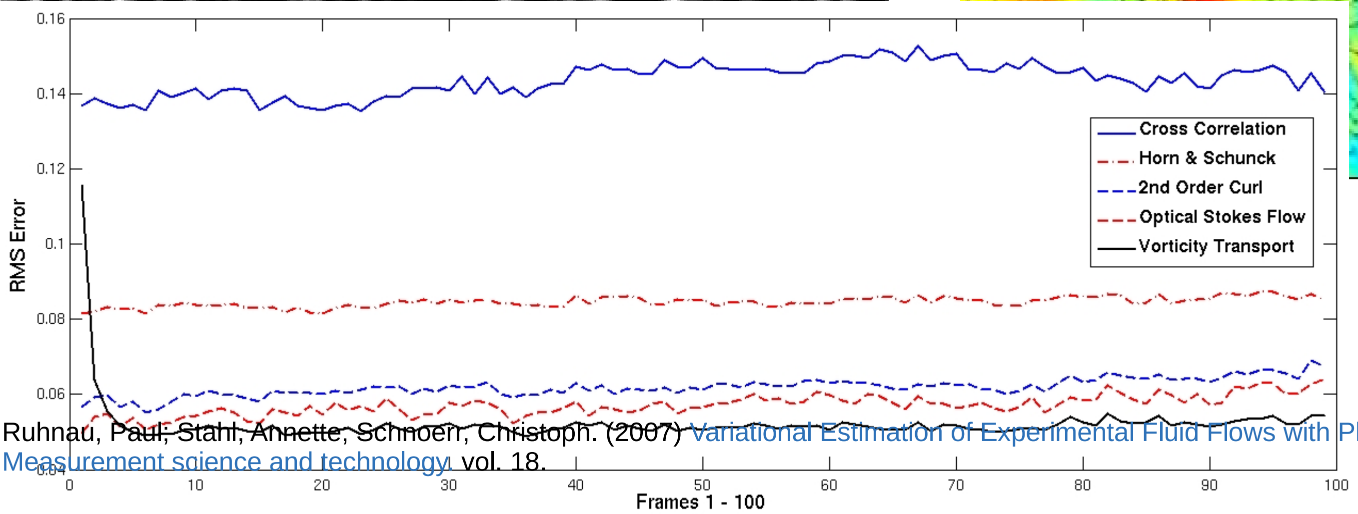
$$E(u) = \frac{1}{2} \int_{\Omega \times [0,T]} \left\{ (\partial_t I + (u \cdot \nabla) I)^2 + \alpha (|\nabla_\theta u_1|^2 + |\nabla_\theta u_2|^2) \right\} dx dt$$

Where ∇_θ is the spatio-temporal gradient $(\partial_{x_1}, \partial_{x_2}, \partial_t)^\top$ of the velocity components.

Variational Optical Flow – Navièr-Stokes

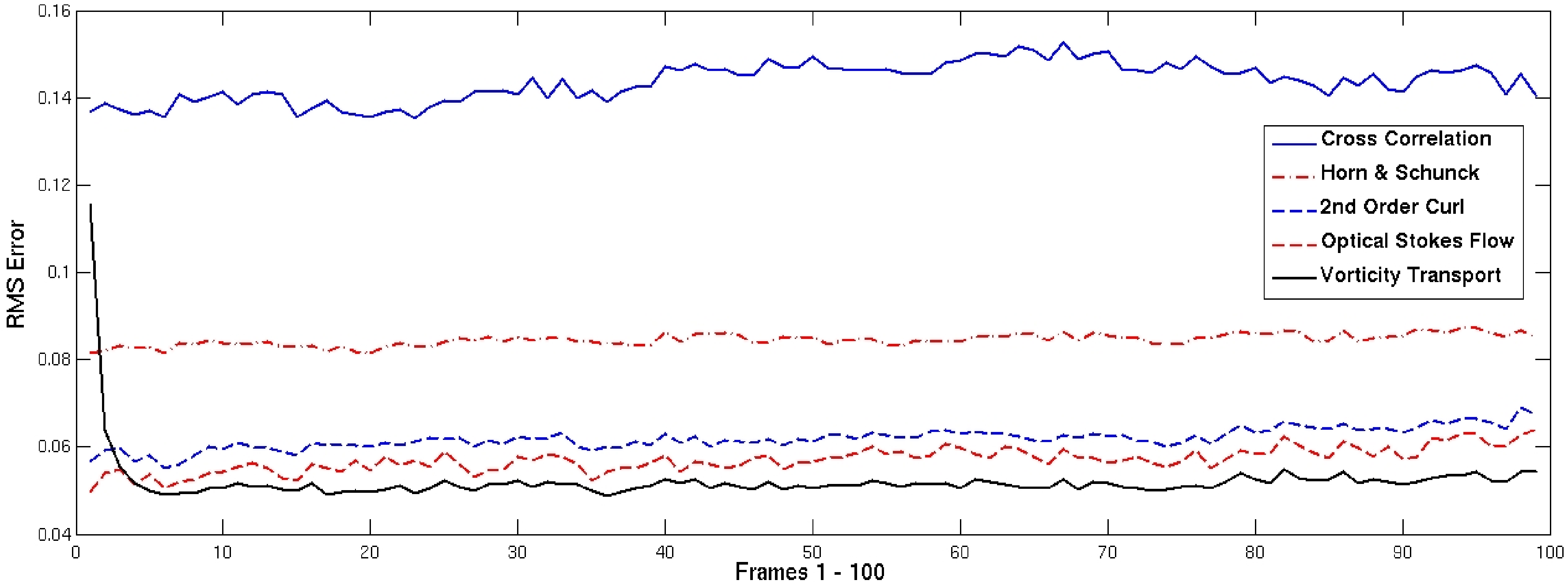






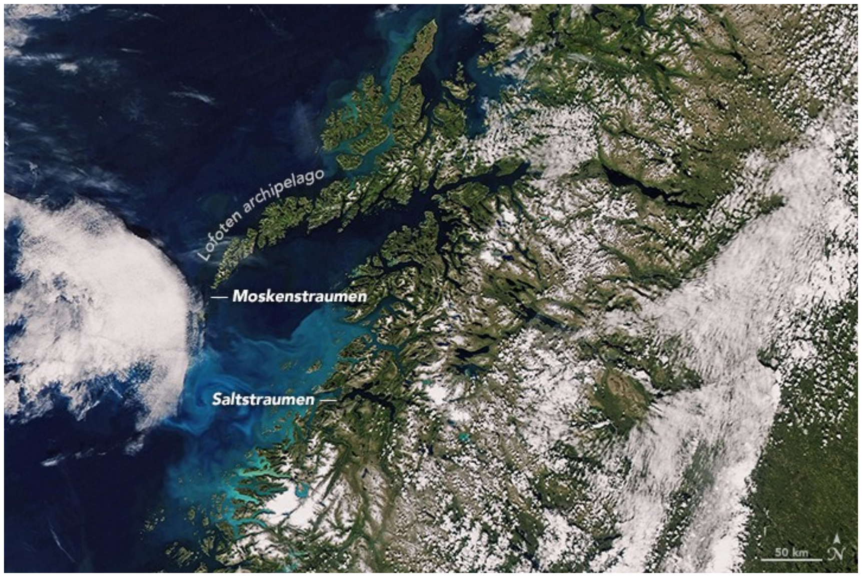
Ruhnau, Paul; Stahl, Annette; Schnoerr, Christoph. (2007) Variational Estimation of Experimental Fluid Flows with Physics-Based Spatio-Temporal Regularization. Measurement science and technology. vol. 18.

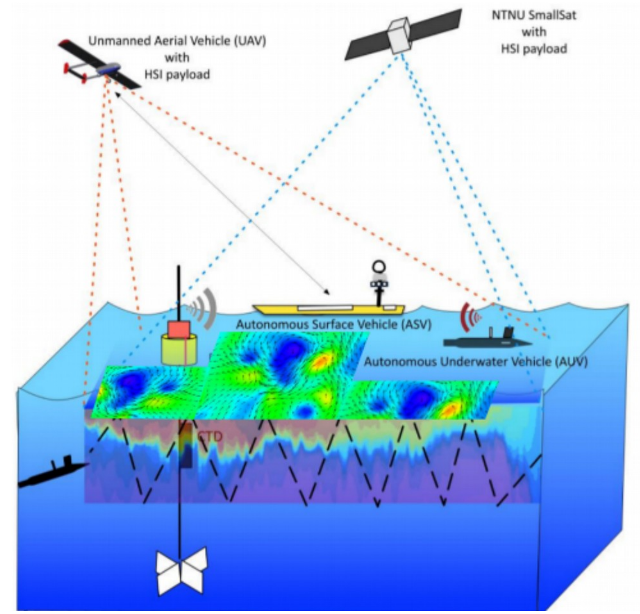
Variational Optical Flow – Navièr-Stokes



Ruhnau, Paul; Stahl, Annette; Schnoerr, Christoph. (2007) Variational Estimation of Experimental Fluid Flows with Physics-Based Spatio-Temporal Regularization. Measurement science and technology. vol. 18.

Ocean Observation





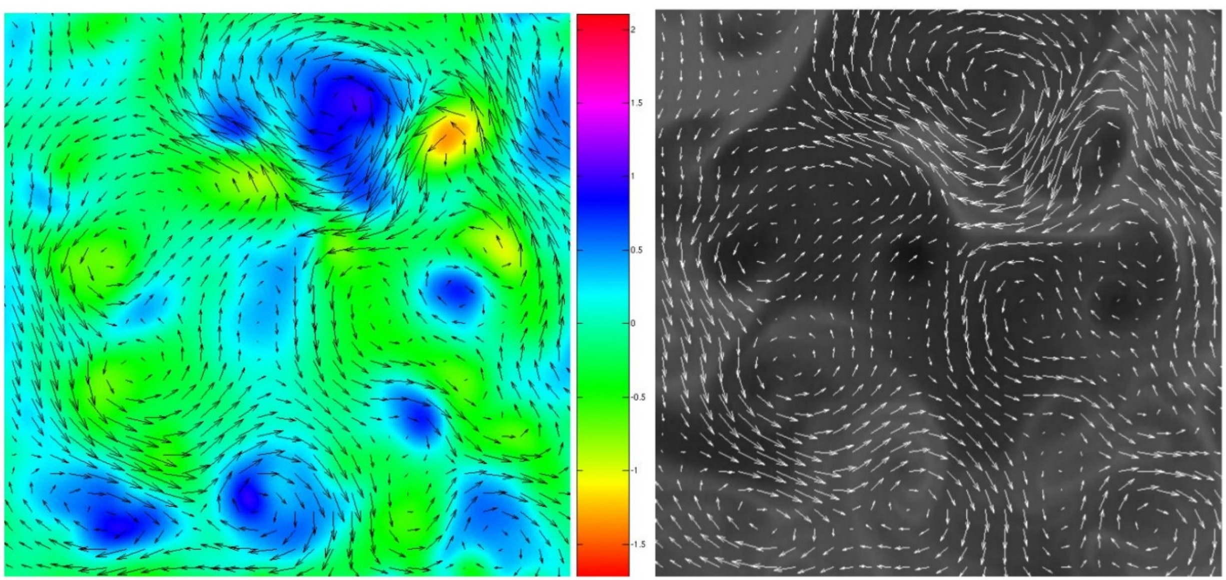


Figure 3. Image based dense velocity field estimates including physical prior knowledge (Navier Stokes Equations NSEs) (**Ruhnau and Stahl 06, 07**).

Summary

Dense Approach – Variational Methods

- analyses at each pixel location
- statements on the existence and uniqueness of solutions
- No multiple processing steps (heuristics)
- modeling assumptions transparent → no hidden assumptions
- fewer tuning parameters
- easily fused

Motion Estimation

- Local: Lucas and Kanade
- Global: Horn and Schunck

Literature

Szeliski Ch 8.4