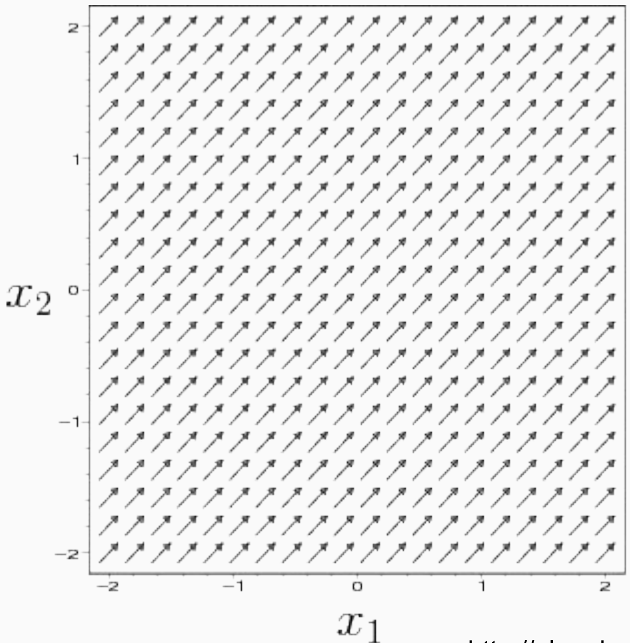
# Order-Adaptive Regularisation for Variational Optical Flow: Global, Local and in Between

Daniel Maurer, Michael Stoll, Andrés Bruhn
presented by David Hägele

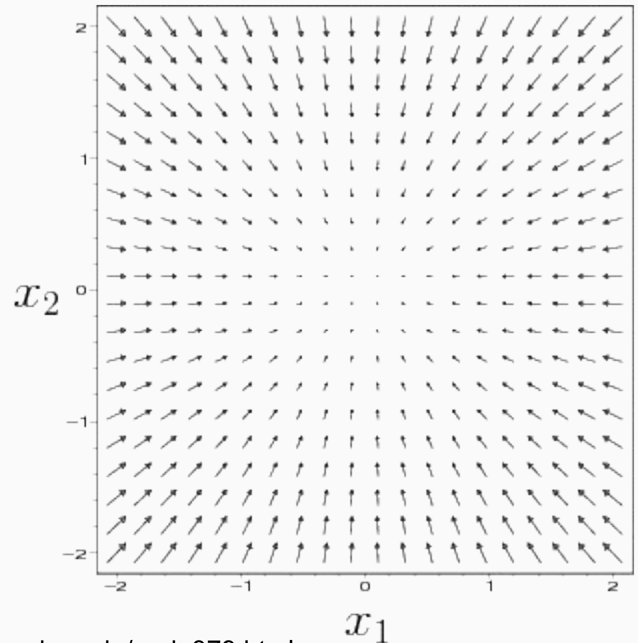
Seminar on Recent Advances in Computer Vision SS18 - University of Stuttgart

# Motivation

Fronto-parallel




Affine




src: http://planning.cs.uiuc.edu/node379.html

$$f(\boldsymbol{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad f(\boldsymbol{x}) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}$$

$$\frac{\partial f}{\partial \boldsymbol{x}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \frac{\partial f}{\partial \boldsymbol{x}} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

# Recap - Variational Optic Flow

Image sequence

$$I_0(\boldsymbol{x}), I_1(\boldsymbol{x}) : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$$

Flow (displacement vector field)

$$\boldsymbol{u} = \begin{pmatrix} u(\boldsymbol{x}) \\ v(\boldsymbol{x}) \end{pmatrix} : \Omega \to \mathbb{R}^2$$

We can find the same pixels in second image when looking in the right place

$$I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) = I_0(\boldsymbol{x})$$

Minimize energy functional to estimate flow

$$E(\boldsymbol{u}) = D(\boldsymbol{u}) + \alpha \cdot R(\boldsymbol{u})$$

# Recap - Data Term and Regularization

Measuring how well I can explain where everything moved (Data Term)

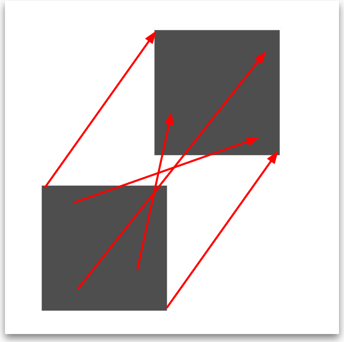
$$\int_{\Omega} p(\boldsymbol{u})\, dx$$

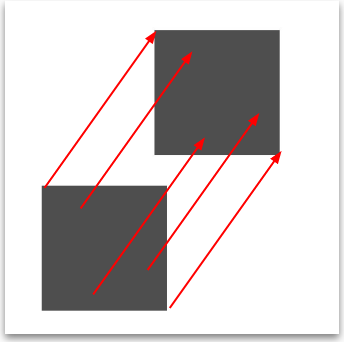
Do pixels have the same values?

$$p(\boldsymbol{u}) = (\ I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_0(\boldsymbol{x})\ )^2$$

Are Edges in the same place?

$$p(\boldsymbol{u}) = ||\ \nabla I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - \nabla I_0(\boldsymbol{x})\ ||_2^2$$





To avoid crazy flow functions, reward sanity (Regularization)

Flow should be smooth!

$$p(\boldsymbol{u}) = ||\ \nabla u\ ||_2^2 + ||\ \nabla v\ ||_2^2$$

# Approach: Data Term

Data Term with two constancy assumptions

$$D(\boldsymbol{u}) = \int_{\Omega} \Psi_D(p_{brightness}(\boldsymbol{u})) + \gamma \cdot \Psi_D(p_{gradient}(\boldsymbol{u}))\; dx$$

Brightness constancy assumption

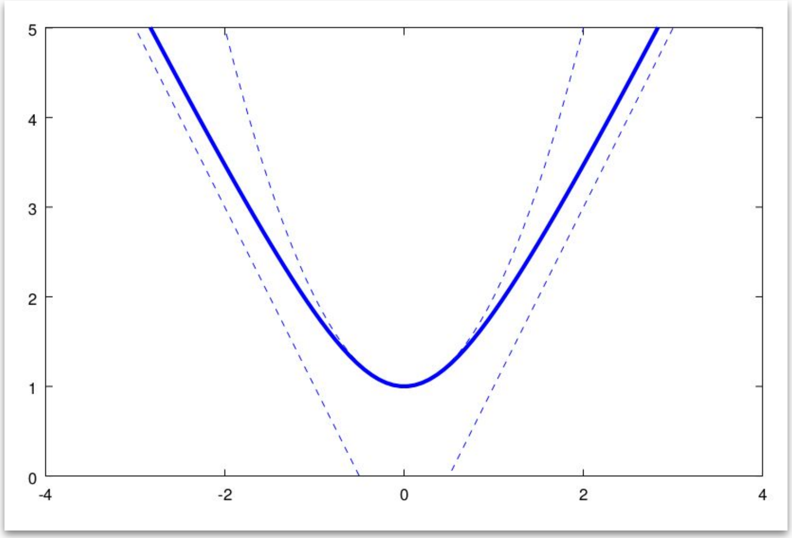
$$p_{brightness}(\boldsymbol{u}) = (\; I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - I_0(\boldsymbol{x})\; )^2$$

Gradient constancy assumption

$$p_{gradient}(\boldsymbol{u}) = ||\; \nabla I_1(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - \nabla I_0(\boldsymbol{x})\; ||_2^2$$

Robustification function
subquadratic Charbonnier penalizer

$$\Psi_D(s^2) = 2\epsilon^2 \sqrt{1 + s^2/\epsilon^2}$$



# Approach: Regularization Terms - Regularization Tensor

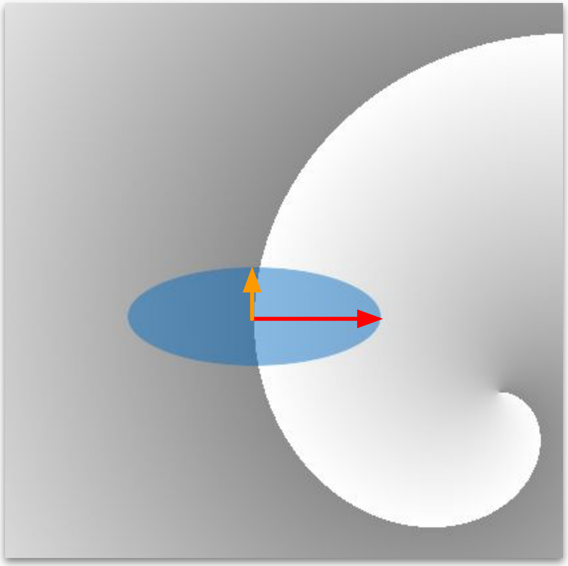
Generalization of the Structure Tensor to arbitrary constancy assumptions

$$S_w(\boldsymbol{x}) = \int w(\tau) \begin{pmatrix} I_x(\boldsymbol{x}-\tau)^2 & I_x(\boldsymbol{x}-\tau) I_y(\boldsymbol{x}-\tau) \\ I_x(\boldsymbol{x}-\tau) I_y(\boldsymbol{x}-\tau) & I_y(\boldsymbol{x}-\tau)^2 \end{pmatrix} d\tau$$

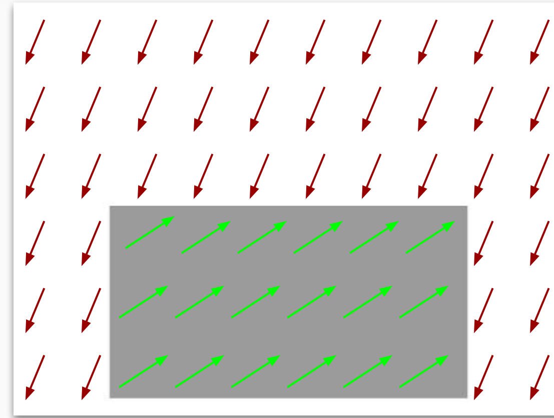
Aggregates the distribution of the gradients within the convolution window

Using eigenvector decomposition we get a summary of this distribution

$$S_w = Q \Lambda Q^\top = \begin{pmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{r}_1^\top \\ \boldsymbol{r}_2^\top \end{pmatrix}$$



# Approach: Regularization Terms - First Order Term

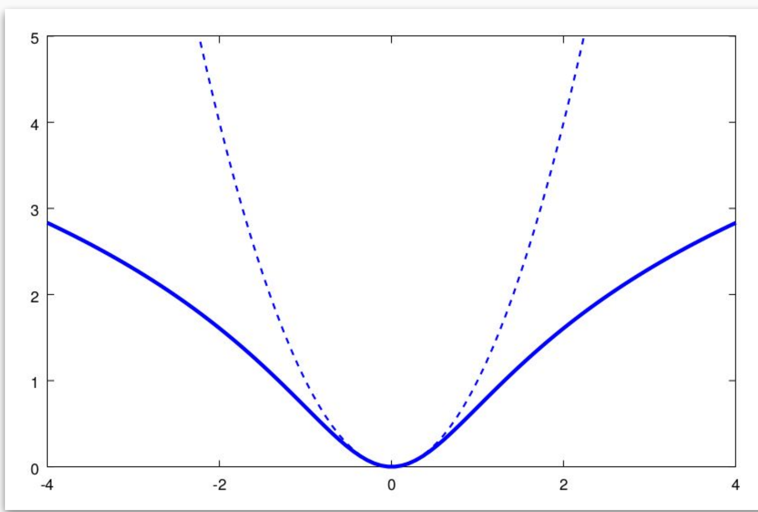


Anisotropic first order regularizer

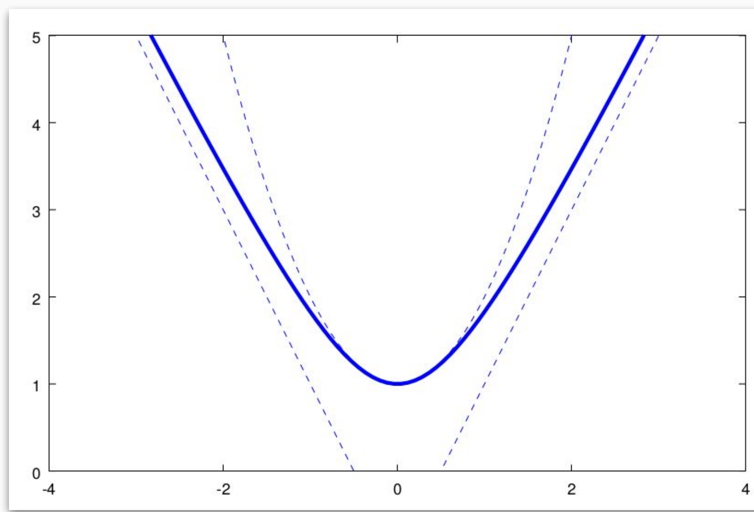
$$R_1(\boldsymbol{u}) = \int_\Omega S_1(\boldsymbol{u})\,dx = \int_\Omega \sum_{l=1}^{2} \Psi_l\Big((\boldsymbol{r}_l^\top \nabla u)^2 + (\boldsymbol{r}_l^\top \nabla v)^2\Big)\,dx$$

Allow flow to change along edges

$$\Psi_1(s^2) = \epsilon^2 \ln(1 + s^2/\epsilon^2)$$




$$\Psi_2(s^2) = 2\epsilon^2 \sqrt{1 + s^2/\epsilon^2}$$



# Approach: Regularization Terms - Second Order Term

Anisotropic second order regularizer

$$R_2(\boldsymbol{u}) = \int_\Omega \inf_{\boldsymbol{a},\boldsymbol{b}} \Big[ S_2(\boldsymbol{u},\boldsymbol{a},\boldsymbol{b}) + \beta \cdot S_{\mathrm{aux}}(\boldsymbol{a},\boldsymbol{b}) \Big] \, dx$$

Flow change should be similar to auxiliary functions

$$S_2(\boldsymbol{u},\boldsymbol{a},\boldsymbol{b}) = \sum_{l=1}^{2} \Psi_l \Big( (\boldsymbol{r}_l^\top (\nabla u - \boldsymbol{a}))^2 + (\boldsymbol{r}_l^\top (\nabla v - \boldsymbol{b})) \Big)$$

Auxiliary functions should be smooth ≈ second derivative of flow should vanish

$$S_{\mathrm{aux}}(\boldsymbol{a},\boldsymbol{b}) = \sum_{l=1}^{2} \Psi_l \Big( \sum_{k=1}^{2} (\boldsymbol{r}_k^\top \mathcal{J}\boldsymbol{a}\, \boldsymbol{r}_l)^2 + (\boldsymbol{r}_k^\top \mathcal{J}\boldsymbol{b}\, \boldsymbol{r}_l)^2 \Big)$$

# Approach: Order Adaption

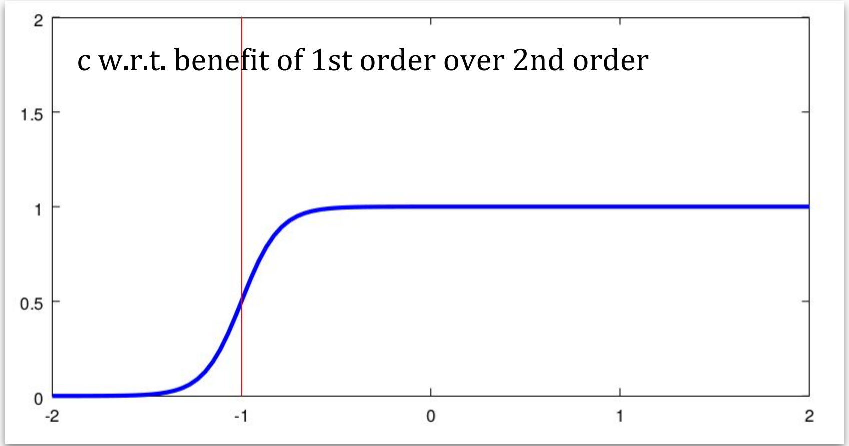
Simple Combination

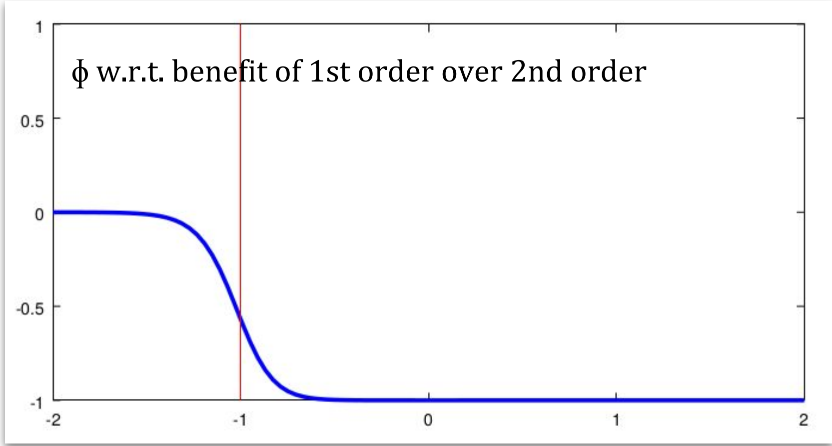
$$R(\boldsymbol{u}) = \int_\Omega \inf_{\boldsymbol{a},\boldsymbol{b}} \Big[ S_1(\boldsymbol{u}) + S_2(\boldsymbol{u},\boldsymbol{a},\boldsymbol{b}) + \beta \cdot S_{\mathrm{aux}}(\boldsymbol{a},\boldsymbol{b}) \Big] \, dx$$

Adaptive Combination (Global Scheme)

$$R(\boldsymbol{u}, c) = \int_\Omega \inf_{\boldsymbol{a},\boldsymbol{b}} \Big[ c \cdot S_1(\boldsymbol{u}) + (1-c) \cdot S_2(\boldsymbol{u},\boldsymbol{a},\boldsymbol{b}) + \beta \cdot S_{\mathrm{aux}}(\boldsymbol{a},\boldsymbol{b}) + \phi_\lambda(c) \Big] \, dx$$

$$c = \frac{1}{1 + e^{-\Delta/\lambda}} \quad \text{with} \quad \Delta = T + \frac{1}{|\Omega|} \int_\Omega S_2 - S_1 \, dx$$



# Approach: Adaption Schemes

Local - different selection for each pixel

$$c_{\text{local}}(\boldsymbol{x}) = \frac{1}{1 + e^{-\Delta/\lambda}} \quad \text{with} \quad \Delta = T + S_2(\boldsymbol{x}) - S_1(\boldsymbol{x})$$

Nonlocal - different selection for each pixel based on its neighborhood

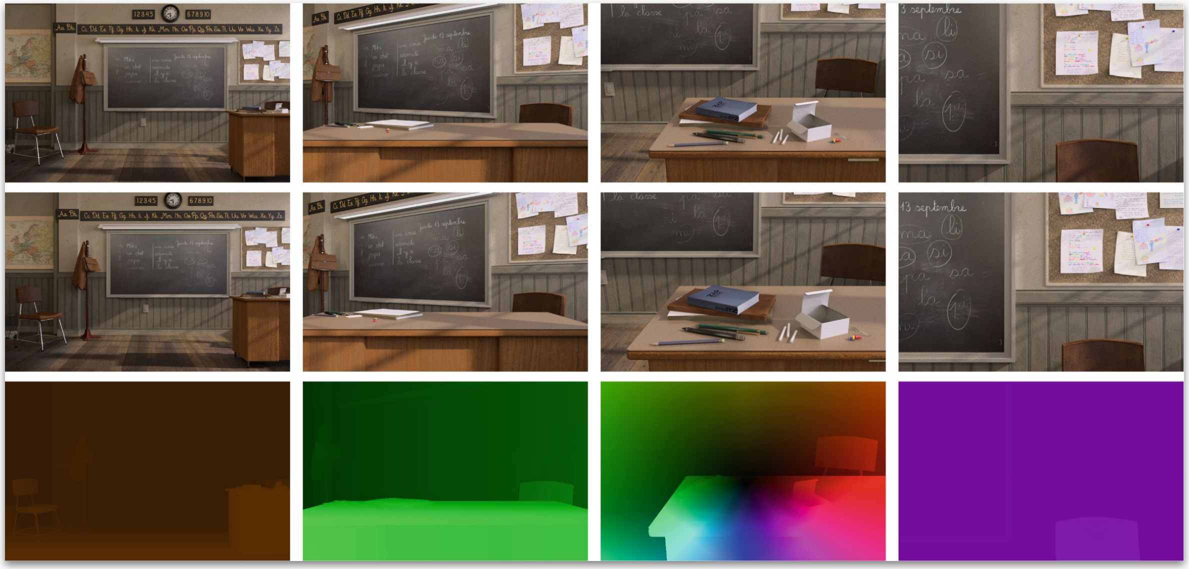
$$c_{\text{nonlocal}}(\boldsymbol{x}) = \frac{1}{1 + e^{-\Delta/\lambda}} \quad \text{with} \quad \Delta = T + \frac{1}{|\mathcal{N}(\boldsymbol{x})|} \int_{\mathcal{N}(x)} S_2 - S_1 \, d\boldsymbol{y}$$

Region-based - different selection per region

$$z : \Omega \to \mathbb{R} \quad \text{(a level-set function)}$$

$$c_{\text{region}}(z) = \frac{1}{1 + e^{-z/\lambda}} \quad \text{and selection term } \Phi_\lambda(c) \text{ replaced by } |\nabla c_{\text{region}}(z)|$$

# Results - Classroom Sequence



| | | Seq. 1 | Seq. 2 | Seq. 3 | Seq. 4 | Avg. | Runtime |
|---|---|---|---|---|---|---|---|
| First order | | 0.129 | 0.358 | 2.038 | **0.088** | 0.653 | 17 s |
| Second order | | 0.141 | 0.370 | 0.669 | 0.102 | 0.321 | 75 s |
| Adaptive order | Global | 0.141 | 0.365 | 0.667 | 0.095 | 0.317 | 100 s |
| Adaptive order | Local | **0.111** | **0.260** | 1.115 | **0.088** | 0.393 | 105 s |
| Adaptive order | Non-local | 0.116 | 0.275 | 0.737 | 0.095 | **0.307** | 120 s |
| Adaptive order | Region | 0.125 | 0.366 | **0.662** | 0.098 | 0.313 | 180 s |

Avg Endpoint Error

# Results - Classroom Sequence

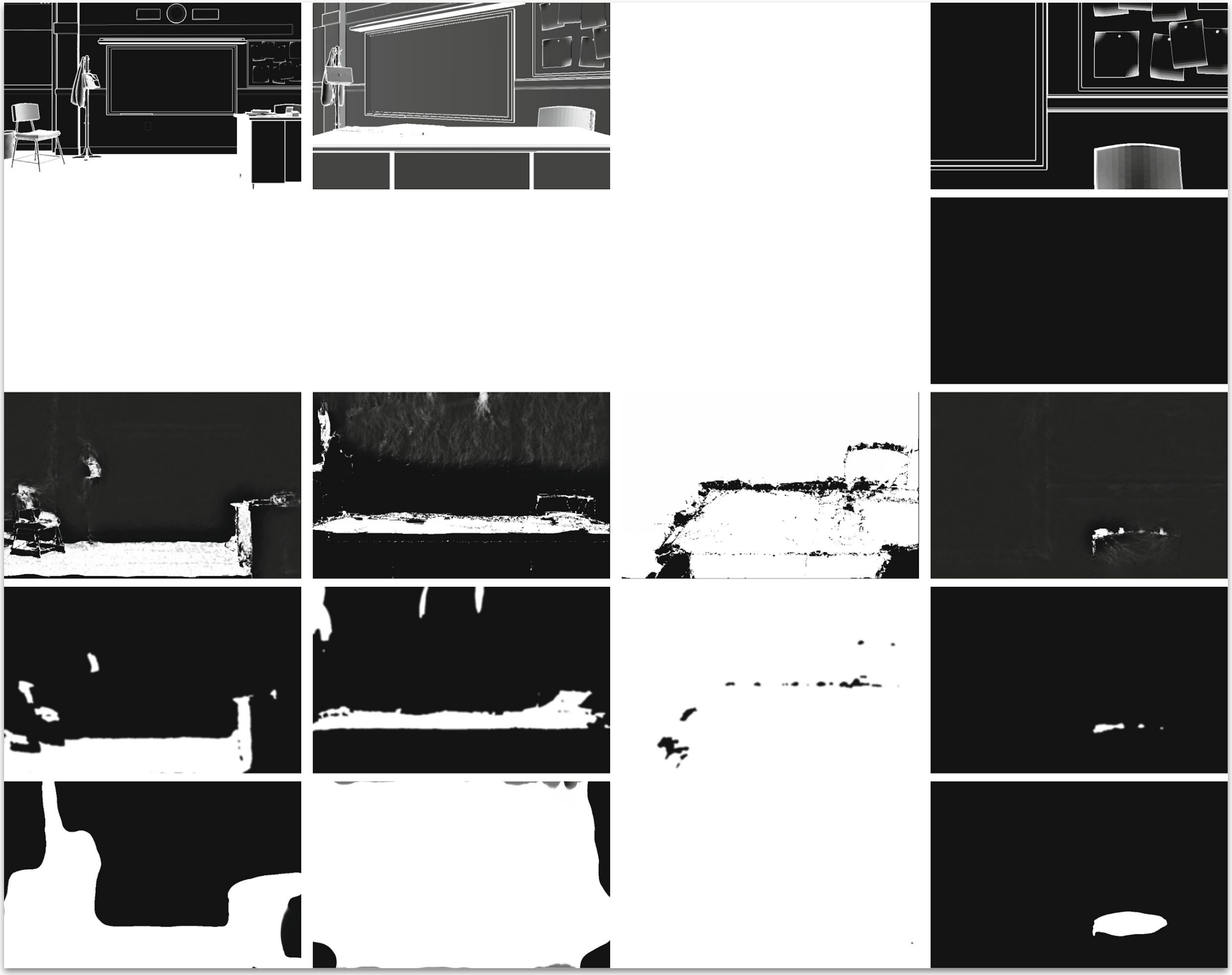
Ground Truth

Global Scheme

Local Scheme

Non-Local Scheme

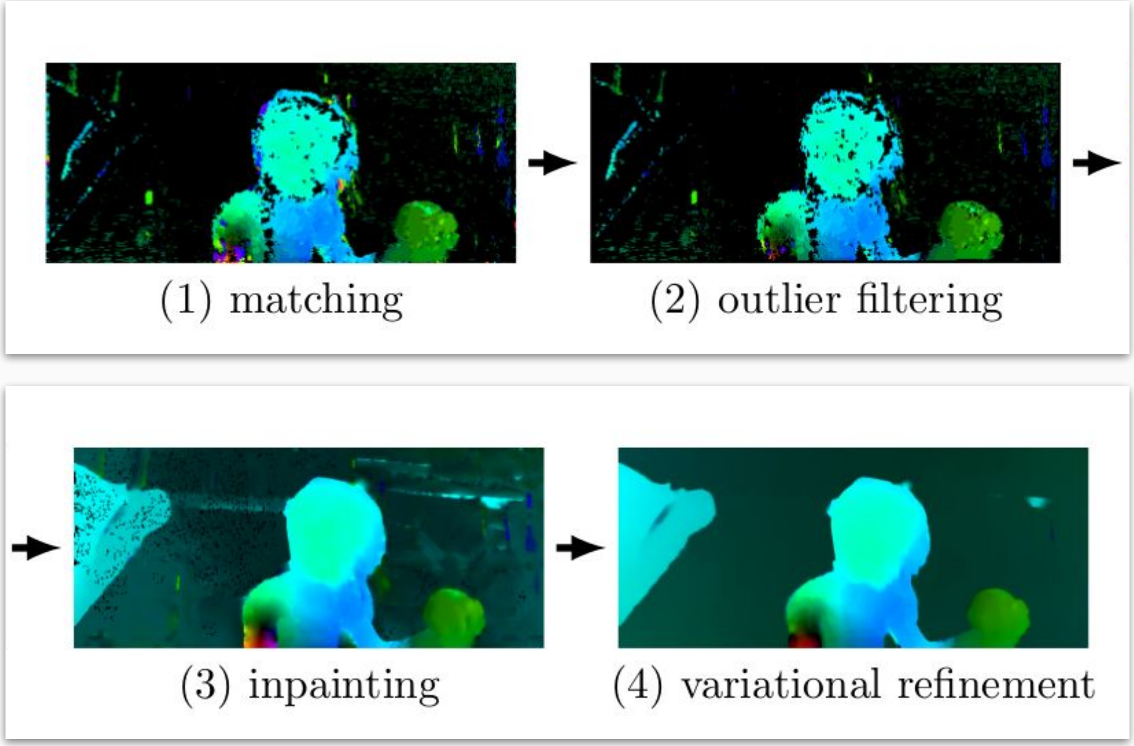
Region-based Scheme



Gradient magnitude of flow field

# Variational Refinement

Pipeline for large displacement optical flow



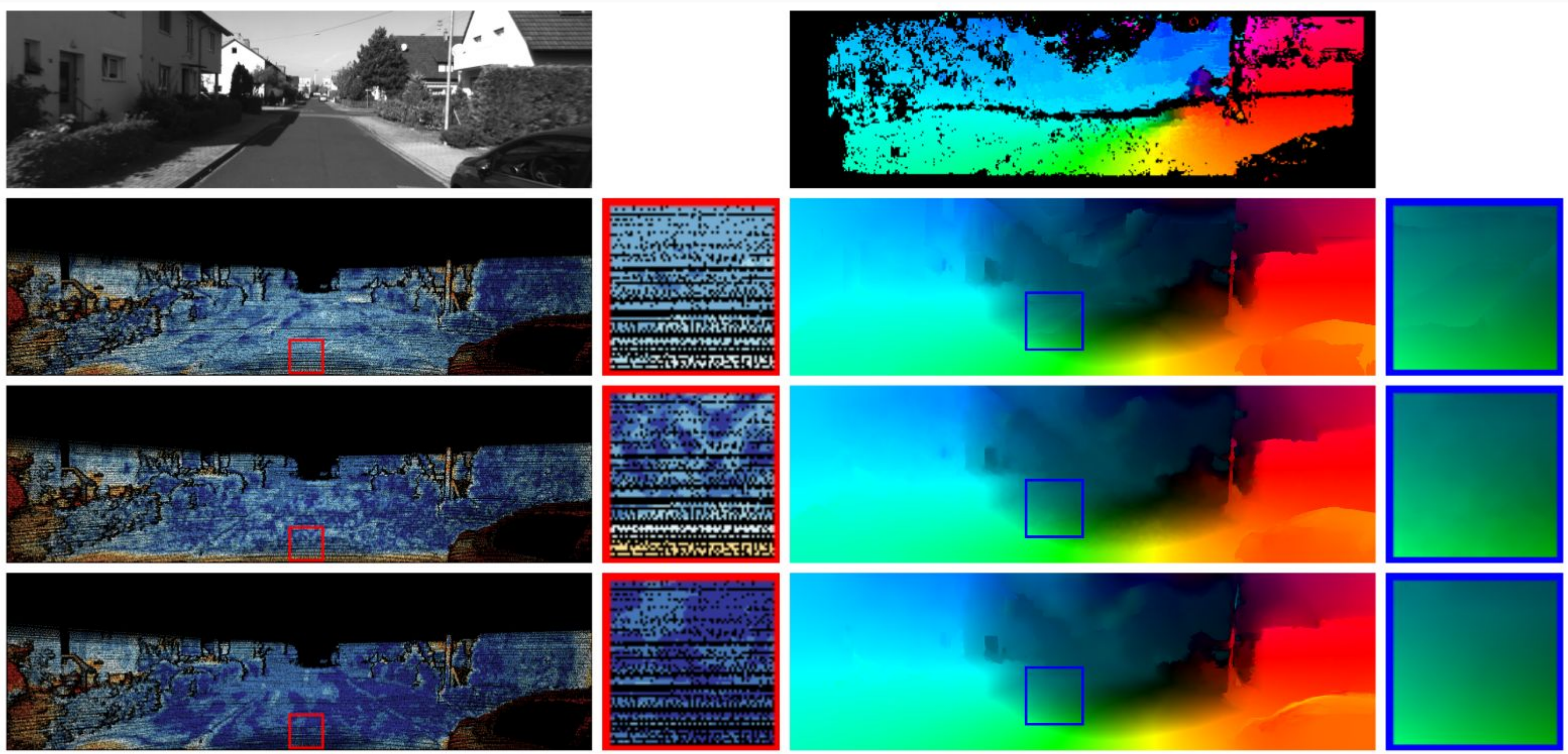

Variational refinement step starts with reasonable initialization of flow

# Variational Refinement - Results

inpainted matches

EpicFlow

OA



Bad Pixels

Flow field

# Variational Refinement - Results

| KITTI 2012 | Out-Noc | Out-All | Avg-Noc | Avg-All |
|---|---|---|---|---|
| ImpPB+SPCI | 4.65 % | 13.47 % | 1.1 px | 2.9 px |
| FlowNet2 | 4.82 % | 8.80 % | 1.0 px | 1.8 px |
| FlowFieldCNN | 4.89 % | 13.01 % | 1.2 px | 3.0 px |
| RicFlow [17] | 4.96 % | 13.04 % | 1.3 px | 3.2 px |
| FlowFields+ | 5.06 % | 13.14 % | 1.2 px | 3.0 px |
| **DF+OIR** | **5.17 %** | **10.43 %** | **1.1 px** | **2.9 px** |
| PatchBatch [11] | 5.29 % | 14.17 % | 1.3 px | 3.3 px |
| SODA-Flow [21] | 5.57 % | 10.71 % | 1.3 px | 2.8 px |
| OAR-Flow [20] | 5.69 % | 10.72 % | 1.4 px | 2.8 px |
| DDF | 5.73 % | 14.18 % | 1.4 px | 3.4 px |
| PH-Flow | 5.76 % | 10.57 % | 1.3 px | 2.9 px |
| FlowFields [1] | 5.77 % | 14.01 % | 1.4 px | 3.5 px |
| CPM-Flow [16] | 5.79 % | 13.70 % | 1.3 px | 3.2 px |
| NLTGV-SC [26] | 5.93 % | 11.96 % | 1.6 px | 3.8 px |
| DDS-DF | 6.03 % | 13.08 % | 1.6 px | 4.2 px |
| TGV2ADCSIFT | 6.20 % | 15.15 % | 1.5 px | 4.5 px |
| S2F-IF | 6.20 % | 15.68 % | 1.4 px | 3.5 px |
| DiscreteFlow [23] | 6.23 % | 16.63 % | 1.3 px | 3.6 px |
| BTF-ILLUM [11] | 6.52 % | 11.03 % | 1.5 px | 2.8 px |
| EpicFlow [28] | 7.88 % | 17.08 % | 1.5 px | 3.8 px |

| KITTI 2015 | Fl-bg | Fl-fg | Fl-all |
|---|---|---|---|
| FlowNet2 | 10.75 % | 8.75 % | 10.41 % |
| DCFlow | 13.10 % | 23.70 % | 14.86 % |
| SOF | 14.63 % | 22.83 % | 15.99 % |
| **DF+OIR** | **15.11 %** | **23.45 %** | **16.50 %** |
| ImpPB+SPCI | 17.25 % | 20.44 % | 17.78 % |
| FlowFieldCNN | 18.33 % | 20.42 % | 18.68 % |
| RicFlow [17] | 18.73 % | 19.09 % | 18.79 % |
| FlowFields+ | 19.51 % | 21.26 % | 19.80 % |
| PatchBatch [11] | 19.98 % | 26.50 % | 21.07 % |
| DDF | 20.36 % | 25.19 % | 21.17 % |
| SODA-Flow [21] | 20.01 % | 29.14 % | 21.53 % |
| DiscreteFlow [23] | 21.53 % | 21.76 % | 21.57 % |
| OAR-Flow [20] | 20.62 % | 27.67 % | 21.79 % |
| CPM-Flow [16] | 22.32 % | 22.81 % | 22.40 % |
| FullFlow [8] | 23.09 % | 24.79 % | 23.37 % |
| SPM-BP | 24.06 % | 24.97 % | 24.21 % |
| EpicFlow [28] | 25.81 % | 28.69 % | 26.29 % |
| DeepFlow [54] | 27.96 % | 31.06 % | 28.48 % |
| HS | 39.90 % | 51.39 % | 41.81 % |
| DB-TV-L1 | 47.52 % | 48.27 % | 47.64 % |

# Summary - End of Presentation