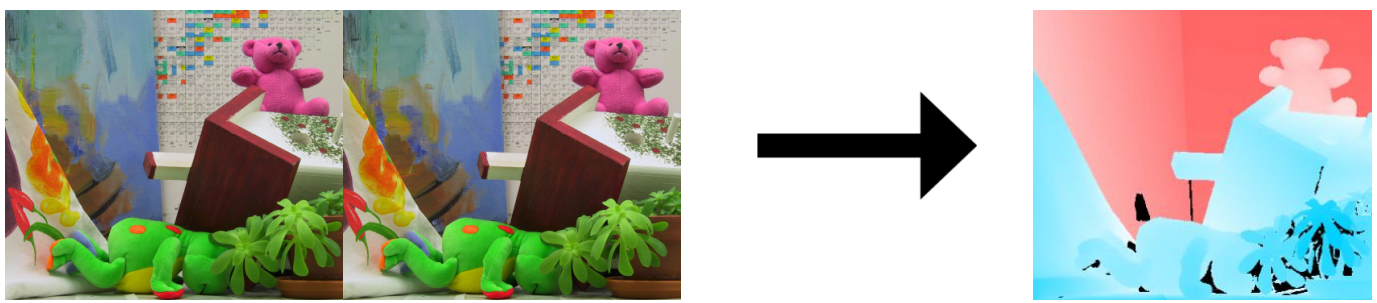
Efficient Nonlocal Regularization for Optical Flow (ECCV 2012)

P. Krähenbühl, V. Koltun (Stanford University)



Lilli Kaufhold, December 18th

Optical Flow

Given: images $I_1, I_2 \in R^N$
Wanted: displacement vector $u \in R^{2N}$ such that pixel $I_1(x)$ corresponds to $I_2(x+u)$

Different Constancy Assumptions:

Grey Value:	$I_2(x+u) - I_1(x) = 0$
Gradient:	$\nabla I_2(x+u) - \nabla I_1(x) = 0$
Hessian:	$\mathcal{H}_2 I_2(x+u) - \mathcal{H}_2 I_1(x) = 0$
Laplacian:	$\Delta I_2(x+u) - \Delta I_1(x) = 0$

Optical Flow

Minimise an energy functional of the form

$$E(u) = E_{Data}(u) + \lambda_L E_L(u)$$

- **data term:** $E_{Data}(u) = \sum_i \|I_2(x_i + u_i) - I_1(x_i)\|^2$
- **smoothness term:** $E_L(u) = \sum_i \sum_{j \in \mathcal{N}_i} \Psi\left((u_i - u_j)^2\right) w_{ij}$
- $\mathbf{\Psi(\cdot)}$: penaliser function $\Psi(x^2) = (x^2 + \epsilon^2)^{0.45}$
- $\mathbf{w_{i,j}}$: weighting function determining the influence pixel j has on pixel i

Content

Nonlocal Regularisation

Energy Functional:

$$E(u) = E_{Data}(u) + \lambda_L E_L(u) + \lambda_N E_N(u)$$

- **nonlocal regulariser:** $E_N(u) = \sum_i \sum_{j \neq i} \Psi\left((u_i - u_j)^2\right) w_{ij}$
- $\mathbf{\Psi(\cdot)}$:penaliser function $\Psi(x^2) = (x^2 + \epsilon^2)^{0.45}$
- $\mathbf{w_{i,j}}$: weighting function determining the influence pixel j has on pixel i

Nonlocal Regularisation

Visualisation of nonlocal weights:

$$w_{i,j} = \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_x^2} - \frac{\|c_i - c_j\|^2}{2\sigma_c^2}\right)$$

Figure: Higher intensity corresponds to higher weight.

Minimisation

$$E(u) = \sum_i \|I_2(x_i + u_i) - I_1(x_i)\|^2 + \lambda_L \sum_i \sum_{j \in \mathcal{N}_i} \Psi\left((u_i - u_j)^2\right) w_{ij}$$
$$+\lambda_N \sum_i \sum_{j \neq i} \Psi\left((u_i - u_j)^2\right) w_{ij}$$

Minimisation Strategy:

- Compute $\nabla E(u)$
- Solve $\nabla E(u) = 0$

- Difficulties:
  - energy functional is not convex nor linear
  - normal minimisation techniques are too slow for the nonlocal term
- E_{data} and E_L easier to handle $\rightarrow$ only focus on E_N

Linearisation

Nonlocal regulariser

$$E_N(u) = \sum_i \sum_{j \neq i} \Psi\left((u_i - u_j)^2\right) w_{ij}$$

- Computation of $\nabla E_N(u)$:

$$\frac{\partial}{\partial u_i} E_N(u) = 2 \sum_{j \neq i} w_{ij} \Psi'\left((u_i - u_j)^2\right)(u_i - u_j)$$

- Linearisation:

$$\frac{\partial}{\partial u_i} E_N(u^{l+1}) = 2 \sum_{j \neq i} w_{ij} \Psi'\left((u_i^l - u_j^l)^2\right)(u_i^{l+1} - u_j^{l+1})$$

Linear System

- Resulting Linear Equation: $(A+B)u^{l+1} = w_B$ with

$$\begin{aligned} Au^{l+1} &= \nabla E_N(u) \\ Bu^{l+1} - w_B &= \nabla E_{data}(u) + \nabla E_L(u) \end{aligned}$$

- The entries of A are given by:

$$\begin{aligned} A_{ij} &= -w_{ij}\Psi'\left((u_i^l - u_j^l)^2\right) \qquad \text{for } i \neq j \\ A_{ii} &= \sum_{j \neq i} w_{ij}\Psi'\left((u_i^l - u_j^l)^2\right) \end{aligned}$$

- Jacobi-Method:

$$u_i^{l+1,k+1} = (A_{ii} + B_{ii})^{-1} \cdot \left(w_{B\,i} - \sum_{j \neq i} (A_{ij} + B_{ij}) u_j^{l+1,k} \right) \qquad \text{for all } i$$

Computational Problems

Problem: Each iteration is very expensive due to the following terms:

$$\begin{aligned} A_{ii} &= \sum_{j \neq i} w_{ij} \Psi' \left((u_i^l - u_j^l)^2 \right) \\ &= \sum_{j \neq i} \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_x^2} - \frac{\|c_i - c_j\|^2}{2\sigma_c^2} \right) \Psi' \left((u_i^l - u_j^l)^2 \right) \\ \sum_{j \neq i} A_{ij} u_j^{l+1,k} &= \sum_{j \neq i} -w_{ij} \Psi' \left((u_i^l - u_j^l)^2 \right) u_j^{l+1,k} \\ &= \sum_{j \neq i} -\exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_x^2} - \frac{\|c_i - c_j\|^2}{2\sigma_c^2} \right) \Psi' \left((u_i^l - u_j^l)^2 \right) u_j^{l+1,k} \end{aligned}$$

Idea: Approximate Ψ with Gaussian kernels and perform an efficient Gaussian convolution.

Reduction to Gaussian Convolution

Approximation with exponential mixtures

$$\Psi(x^2) \approx \mu_{\omega,\sigma}(x^2) := T - \sum_{n=1}^{K} \omega_n \exp(-\frac{x^2}{2\sigma_n^2})$$

$$\Rightarrow \Psi'(x^2) \approx \sum_{n=1}^{K} \frac{\omega_n}{2\sigma_n^2} \exp(-\frac{x^2}{2\sigma_n^2})$$

Minimise

$$\int_{-\infty}^{\infty} \left(\mu_{\omega,\sigma}(x^2) - \tilde{\Psi}(x^2)\right)^2 dx$$

with truncated penalty function $\tilde{\Psi}$
in order to optimise the parameters ω and σ.

Influence of Parameters: K

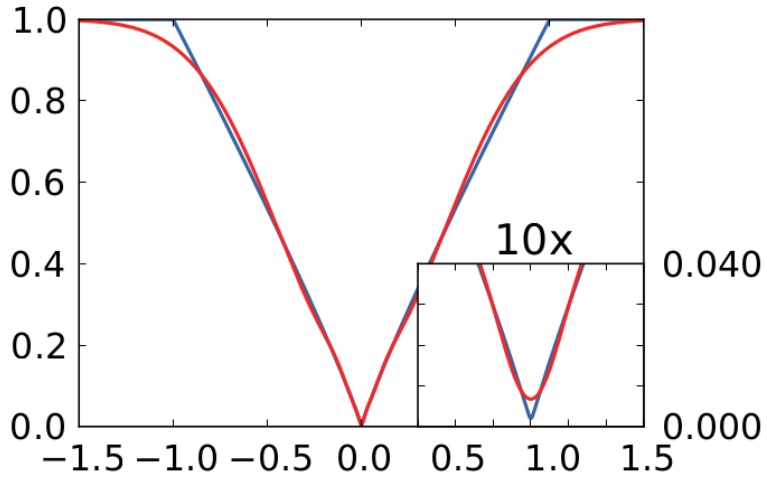


Figure: Penalty function $\Psi(x^2) = (x^2 + \epsilon^2)^{0.45}$ and its Approximation

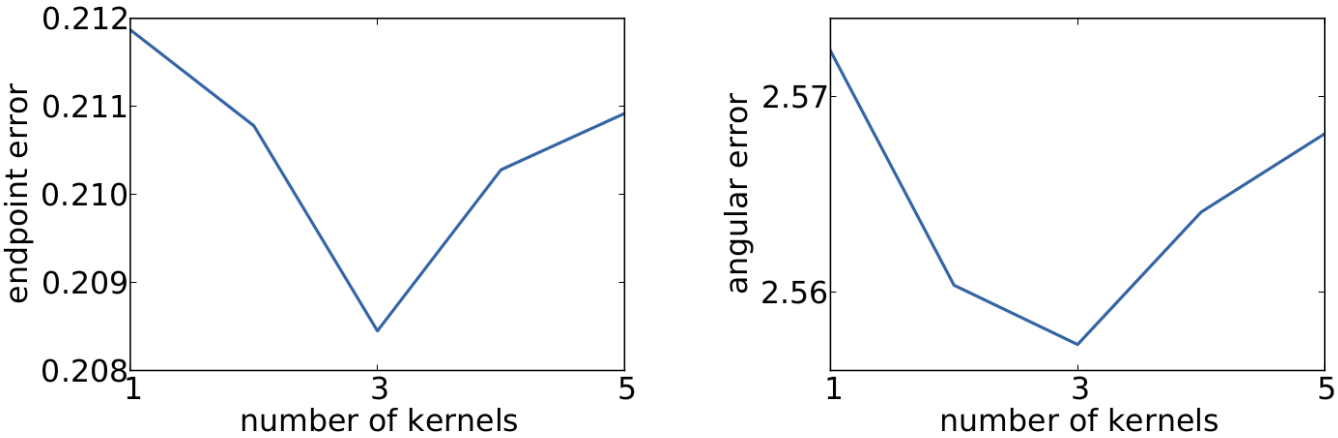


Figure: Average endpoint and angular error for varying number of kernels

Influence of Parameters: σ_c, σ_x and λ_N

$$E_N(u) = \lambda_N \sum_i \sum_{j \neq i} \Psi\left((u_i - u_j)^2\right) \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_x^2} - \frac{\|c_i - c_j\|^2}{2\sigma_c^2}\right)$$

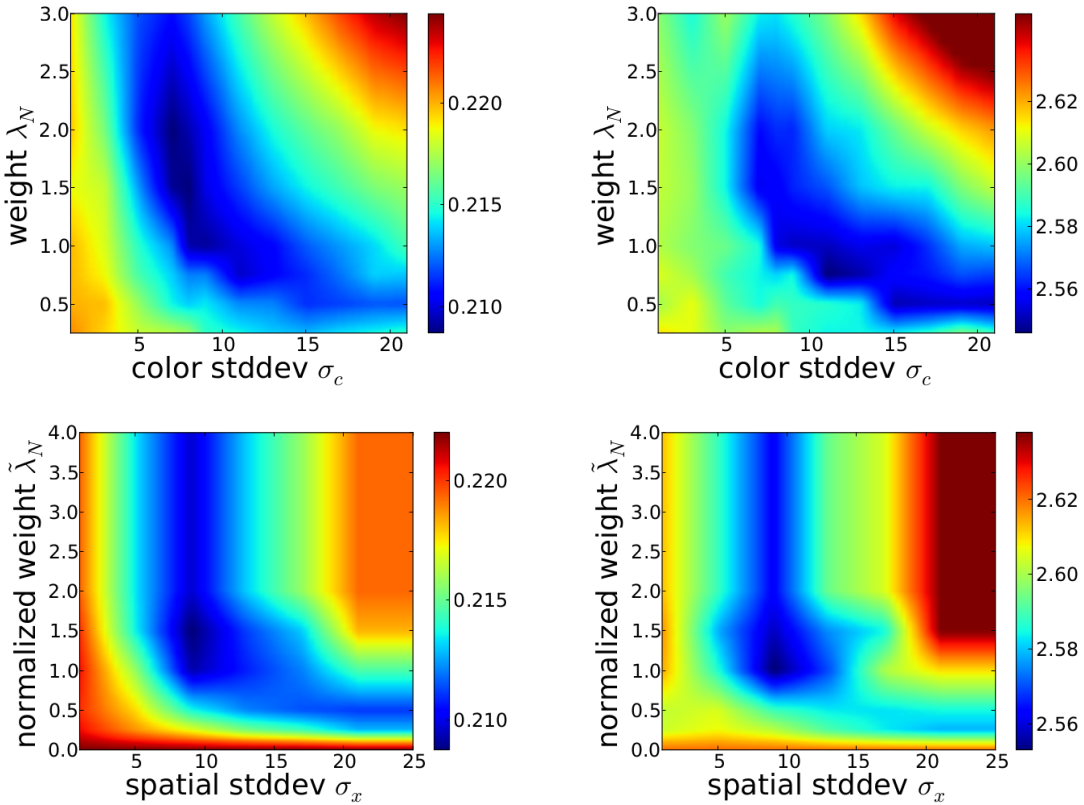


Figure: **top row:** endpoint and angular errors for varying σ_c and λ_N
bottom row: endpoint and angular errors for varying σ_x and λ_N

Results

Figure: **top left:** first image frame **top right:** ground truth flow **bottom left:** result without nonlocal regularisation **bottom right:** result with nonlocal regularisation

Results

Average angle error	avg. rank	Army (Hidden texture) GT all	Army (Hidden texture) GT disc	Army (Hidden texture) GT untext	Army (Hidden texture) im0 all	Army (Hidden texture) im0 disc	Army (Hidden texture) im0 untext	Army (Hidden texture) im1 all	Army (Hidden texture) im1 disc	Army (Hidden texture) im1 untext	Mequon (Hidden texture) GT all	Mequon (Hidden texture) GT disc	Mequon (Hidden texture) GT untext	Mequon (Hidden texture) im0 all	Mequon (Hidden texture) im0 disc	Mequon (Hidden texture) im0 untext	Mequon (Hidden texture) im1 all	Mequon (Hidden texture) im1 disc	Mequon (Hidden texture) im1 untext
NN-field [73]	4.9	2.89 3	8.13 8	**2.11** 1	2.10 3	7.15 4	1.77 6	2.35 2	3.05 3	**1.60** 1	**1.89** 1	**5.20** 1	**1.37** 1	2.43 26	3.70 26	1.95 20	**1.01** 1	**2.25** 1	**0.53** 1
nLayers [57]	8.3	**2.80** 1	**7.42** 1	2.20 3	2.71 15	7.24 5	2.55 33	**2.30** 1	**3.02** 1	1.70 2	2.62 5	6.95 2	2.09 3	2.29 19	3.46 13	1.89 17	1.38 8	3.06 11	1.29 8
MDP-Flow2 [70]	9.3	3.23 19	7.93 5	2.60 12	**1.92** 1	**6.64** 1	**1.52** 1	2.77 12	3.50 9	2.16 12	2.86 7	8.58 7	2.70 14	2.00 9	3.50 18	1.59 8	1.28 5	2.67 5	0.89 3
ADF [67]	11.9	2.98 6	8.32 11	2.28 4	2.27 5	8.35 11	1.81 7	2.64 9	3.55 10	1.81 4	3.02 8	9.08 9	2.38 6	2.29 19	3.48 15	2.07 23	1.34 6	3.03 9	1.11 4
FC-Layers-FF [77]	12.9	3.02 8	7.87 4	2.61 13	2.72 16	9.35 18	2.29 21	2.49 4	3.19 4	2.03 10	3.39 19	8.92 8	2.83 21	2.83 41	3.92 37	2.80 39	1.25 4	2.57 4	1.20 6
Layers++ [37]	14.0	3.11 9	8.22 9	2.79 23	2.43 9	7.02 3	2.24 17	2.35 2	**3.02** 1	1.96 6	3.81 27	11.4 23	3.22 31	2.74 37	4.01 42	2.35 29	1.45 10	3.05 10	1.79 18
Efficient-NL [60]	14.6	2.99 7	8.23 10	2.28 4	2.72 16	8.95 15	2.25 19	2.61 7	3.48 8	1.96 6	3.31 15	8.33 5	2.59 10	2.60 31	3.75 27	2.54 35	1.60 14	3.02 8	1.66 13
LME [72]	14.8	3.15 13	8.04 7	2.31 7	1.95 2	6.65 2	1.59 3	2.85 18	3.61 13	2.42 23	3.47 22	12.8 28	3.17 28	2.12 13	3.53 21	1.73 11	1.34 6	2.75 6	1.18 5
ALD-Flow [68]	15.3	2.82 2	7.86 3	2.16 2	2.84 21	10.1 22	1.86 10	2.69 10	3.60 12	1.85 5	2.79 6	11.3 22	2.32 5	2.07 10	3.25 5	3.10 54	2.03 25	5.11 26	1.94 20
FESL [75]	15.4	2.96 5	7.70 2	2.54 10	3.26 37	10.4 24	2.56 34	2.57 6	3.40 7	2.12 11	2.60 4	7.65 3	2.30 4	2.64 35	4.22 47	2.47 32	1.75 18	3.49 16	1.71 15
IROF++ [58]	15.9	3.17 15	8.69 15	2.61 13	2.79 18	9.61 19	2.33 22	2.74 11	3.57 11	2.19 13	3.20 13	9.70 15	2.71 15	1.96 8	3.45 12	1.22 5	1.80 19	4.06 20	2.50 28
SCR [74]	16.2	3.12 10	8.48 12	2.59 11	2.95 27	10.4 24	2.35 23	2.81 14	3.64 14	2.30 15	3.02 8	8.29 4	2.39 8	2.77 40	3.79 30	2.89 47	1.39 9	2.85 7	1.60 12
Sparse-NonSparse [56]	17.4	3.14 12	8.75 17	2.76 22	3.02 30	10.6 27	2.43 27	2.85 18	3.75 19	2.33 17	3.28 14	9.40 12	2.73 16	2.42 25	3.31 6	2.69 37	1.47 11	3.07 12	1.66 13
TC-Flow [46]	17.7	2.91 4	8.00 6	2.34 8	2.18 4	8.77 12	**1.52** 1	2.78 13	3.73 18	1.96 6	3.08 10	11.4 23	2.66 11	1.94 6	3.43 11	3.20 58	3.06 31	7.04 30	4.08 51
LSM [39]	18.7	3.12 10	8.62 14	2.75 21	3.00 29	10.5 26	2.44 28	2.82 15	3.68 15	2.36 18	3.38 18	9.41 13	2.81 19	2.69 36	3.52 19	2.84 42	1.59 13	3.38 15	1.80 19
Ramp [62]	19.3	3.18 17	8.83 18	2.73 20	2.89 24	10.1 22	2.44 28	2.82 15	3.69 17	2.29 14	3.37 17	9.31 11	2.93 22	2.62 33	3.38 10	3.19 57	1.54 12	3.21 13	2.24 24
COFM [59]	20.0	3.17 15	9.90 33	2.46 9	2.41 8	8.34 10	1.92 11	3.08 24	3.92 23	3.25 47	3.83 28	10.9 18	3.15 27	2.20 17	3.35 7	2.91 50	1.62 16	2.56 3	2.09 21
PMF [76]	20.1	3.61 28	9.07 20	2.62 15	2.40 6	8.05 7	1.83 8	2.54 5	3.27 5	1.71 3	3.59 25	11.1 21	3.46 34	4.07 64	6.18 71	4.02 66	1.06 2	2.38 2	1.25 7
Classic+NL [31]	21.6	3.20 18	8.72 16	2.81 24	3.02 30	10.6 27	2.44 28	2.83 17	3.68 15	2.31 16	3.40 20	9.09 10	2.76 18	2.87 43	3.82 32	2.86 45	1.67 17	3.53 17	2.26 26
TV-L1-MCT [64]	21.6	3.16 14	8.48 12	2.71 19	3.28 38	10.8 34	2.60 38	2.94 21	3.79 20	2.63 32	3.50 23	9.75 16	3.06 25	2.08 11	3.35 7	2.29 27	1.95 22	3.89 19	2.71 31
SimpleFlow [49]	23.7	3.35 21	9.20 23	2.98 28	3.18 34	10.7 31	2.71 40	2.91 20	3.79 20	2.47 24	3.59 25	9.49 14	2.99 23	2.39 23	3.46 13	2.24 26	1.60 14	3.56 18	1.57 10
Direct ZNCC [66]	24.7	3.50 26	8.96 19	2.70 18	2.46 11	9.21 17	1.83 8	3.32 33	4.34 34	2.60 29	4.60 42	14.6 39	4.31 50	2.62 33	3.64 25	3.09 53	1.96 23	4.70 23	1.58 11

Average angle error	avg. rank	Grove (Synthetic) GT all	Grove (Synthetic) GT disc	Grove (Synthetic) GT untext	Grove (Synthetic) im0 all	Grove (Synthetic) im0 disc	Grove (Synthetic) im0 untext	Grove (Synthetic) im1 all	Grove (Synthetic) im1 disc	Grove (Synthetic) im1 untext	Urban (Synthetic) GT all	Urban (Synthetic) GT disc	Urban (Synthetic) GT untext	Urban (Synthetic) im0 all	Urban (Synthetic) im0 disc	Urban (Synthetic) im0 untext	Urban (Synthetic) im1 all	Urban (Synthetic) im1 disc	Urban (Synthetic) im1 untext
NN-field [73]	4.9	2.35 2	3.05 3	**1.60** 1	**1.89** 1	**5.20** 1	**1.37** 1	2.43 26	3.70 26	1.95 20	**1.01** 1	**2.25** 1	**0.53** 1	2.10 3	7.15 4	1.77 6	2.35 2	3.05 3	**1.60** 1
nLayers [57]	8.3	**2.30** 1	**3.02** 1	1.70 2	2.62 5	6.95 2	2.09 3	2.29 19	3.46 13	1.89 17	1.38 8	3.06 11	1.29 8	2.71 15	7.24 5	2.55 33	**2.30** 1	**3.02** 1	1.70 2
MDP-Flow2 [70]	9.3	2.77 12	3.50 9	2.16 12	2.86 7	8.58 7	2.70 14	2.00 9	3.50 18	1.59 8	1.28 5	2.67 5	0.89 3	**1.92** 1	**6.64** 1	**1.52** 1	2.77 12	3.50 9	2.16 12
ADF [67]	11.9	2.64 9	3.55 10	1.81 4	3.02 8	9.08 9	2.38 6	2.29 19	3.48 15	2.07 23	1.34 6	3.03 9	1.11 4	2.27 5	8.35 11	1.81 7	2.64 9	3.55 10	1.81 4
FC-Layers-FF [77]	12.9	2.49 4	3.19 4	2.03 10	3.39 19	8.92 8	2.83 21	2.83 41	3.92 37	2.80 39	1.25 4	2.57 4	1.20 6	2.72 16	9.35 18	2.29 21	2.49 4	3.19 4	2.03 10
Layers++ [37]	14.0	2.35 2	**3.02** 1	1.96 6	3.81 27	11.4 23	3.22 31	2.74 37	4.01 42	2.35 29	1.45 10	3.05 10	1.79 18	2.43 9	7.02 3	2.24 17	2.35 2	**3.02** 1	1.96 6
Efficient-NL [60]	14.6	2.61 7	3.48 8	1.96 6	3.31 15	8.33 5	2.59 10	2.60 31	3.75 27	2.54 35	1.60 14	3.02 8	1.66 13	2.72 16	8.95 15	2.25 19	2.61 7	3.48 8	1.96 6
LME [72]	14.8	2.85 18	3.61 13	2.42 23	3.47 22	12.8 28	3.17 28	2.12 13	3.53 21	1.73 11	1.34 6	2.75 6	1.18 5	1.95 2	6.65 2	1.59 3	2.85 18	3.61 13	2.42 23
ALD-Flow [68]	15.3	2.69 10	3.60 12	1.85 5	2.79 6	11.3 22	2.32 5	2.07 10	3.25 5	3.10 54	2.03 25	5.11 26	1.94 20	2.84 21	10.1 22	1.86 10	2.69 10	3.60 12	1.85 5
FESL [75]	15.4	2.57 6	3.40 7	2.12 11	2.60 4	7.65 3	2.30 4	2.64 35	4.22 47	2.47 32	1.75 18	3.49 16	1.71 15	3.26 37	10.4 24	2.56 34	2.57 6	3.40 7	2.12 11
IROF++ [58]	15.9	2.74 11	3.57 11	2.19 13	3.20 13	9.70 15	2.71 15	1.96 8	3.45 12	1.22 5	1.80 19	4.06 20	2.50 28	2.79 18	9.61 19	2.33 22	2.74 11	3.57 11	2.19 13
SCR [74]	16.2	2.81 14	3.64 14	2.30 15	3.02 8	8.29 4	2.39 8	2.77 40	3.79 30	2.89 47	1.39 9	2.85 7	1.60 12	2.95 27	10.4 24	2.35 23	2.81 14	3.64 14	2.30 15
Sparse-NonSparse [56]	17.4	2.85 18	3.75 19	2.33 17	3.28 14	9.40 12	2.73 16	2.42 25	3.31 6	2.69 37	1.47 11	3.07 12	1.66 13	3.02 30	10.6 27	2.43 27	2.85 18	3.75 19	2.33 17
TC-Flow [46]	17.7	2.78 13	3.73 18	1.96 6	3.08 10	11.4 23	2.66 11	1.94 6	3.43 11	3.20 58	3.06 31	7.04 30	4.08 51	2.18 4	8.77 12	**1.52** 1	2.78 13	3.73 18	1.96 6
LSM [39]	18.7	2.82 15	3.68 15	2.36 18	3.38 18	9.41 13	2.81 19	2.69 36	3.52 19	2.84 42	1.59 13	3.38 15	1.80 19	3.00 29	10.5 26	2.44 28	2.82 15	3.68 15	2.36 18
Ramp [62]	19.3	2.82 15	3.69 17	2.29 14	3.37 17	9.31 11	2.93 22	2.62 33	3.38 10	3.19 57	1.54 12	3.21 13	2.24 24	2.89 24	10.1 22	2.44 28	2.82 15	3.69 17	2.29 14
COFM [59]	20.0	3.08 24	3.92 23	3.25 47	3.83 28	10.9 18	3.15 27	2.20 17	3.35 7	2.91 50	1.62 16	2.56 3	2.09 21	2.41 8	8.34 10	1.92 11	3.08 24	3.92 23	3.25 47
PMF [76]	20.1	2.54 5	3.27 5	1.71 3	3.59 25	11.1 21	3.46 34	4.07 64	6.18 71	4.02 66	1.06 2	2.38 2	1.25 7	2.40 6	8.05 7	1.83 8	2.54 5	3.27 5	1.71 3
Classic+NL [31]	21.6	2.83 17	3.68 15	2.31 16	3.40 20	9.09 10	2.76 18	2.87 43	3.82 32	2.86 45	1.67 17	3.53 17	2.26 26	3.02 30	10.6 27	2.44 28	2.83 17	3.68 15	2.31 16
TV-L1-MCT [64]	21.6	2.94 21	3.79 20	2.63 32	3.50 23	9.75 16	3.06 25	2.08 11	3.35 7	2.29 27	1.95 22	3.89 19	2.71 31	3.28 38	10.8 34	2.60 38	2.94 21	3.79 20	2.63 32
SimpleFlow [49]	23.7	2.91 20	3.79 20	2.47 24	3.59 25	9.49 14	2.99 23	2.39 23	3.46 13	2.24 26	1.60 14	3.56 18	1.57 10	3.18 34	10.7 31	2.71 40	2.91 20	3.79 20	2.47 24
Direct ZNCC [66]	24.7	3.32 33	4.34 34	2.60 29	4.60 42	14.6 39	4.31 50	2.62 33	3.64 25	3.09 53	1.96 23	4.70 23	1.58 11	2.46 11	9.21 17	1.83 8	3.32 33	4.34 34	2.60 29

Figure: Screenshot of the current Middlebury benchmark ranking [http://vision.middlebury.edu/flow/eval/results/results-a1.php]

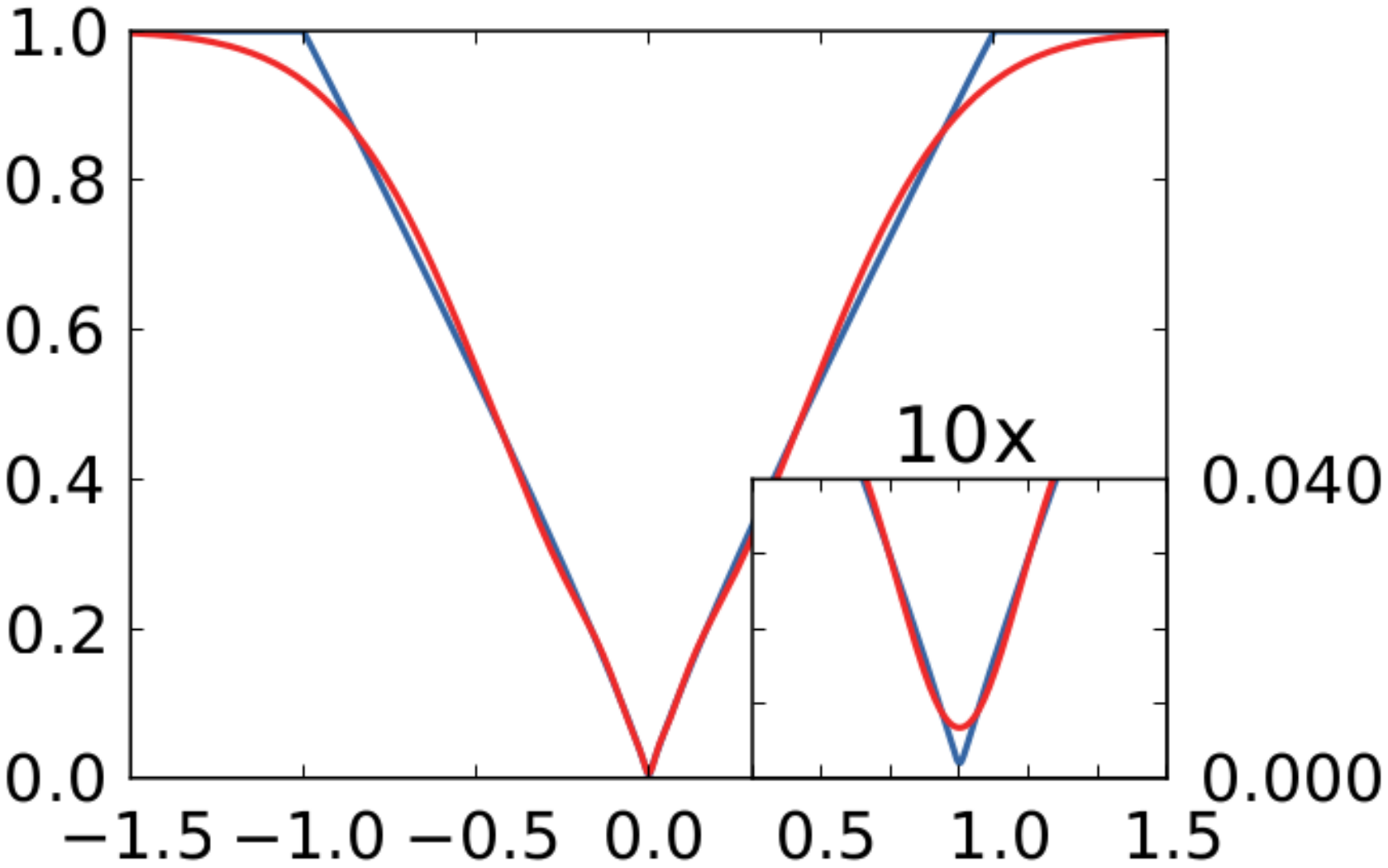
Summary

- Use a nonlocal smoothness term in addition to the commonly used local one
- Approximate the penalty function with Gaussian kernels and use a fast Gaussian convolution method
- Can be computed very efficiently
- Nonlocality improves the results of this optical flow method
- Can maybe even improve the best method there currently is?

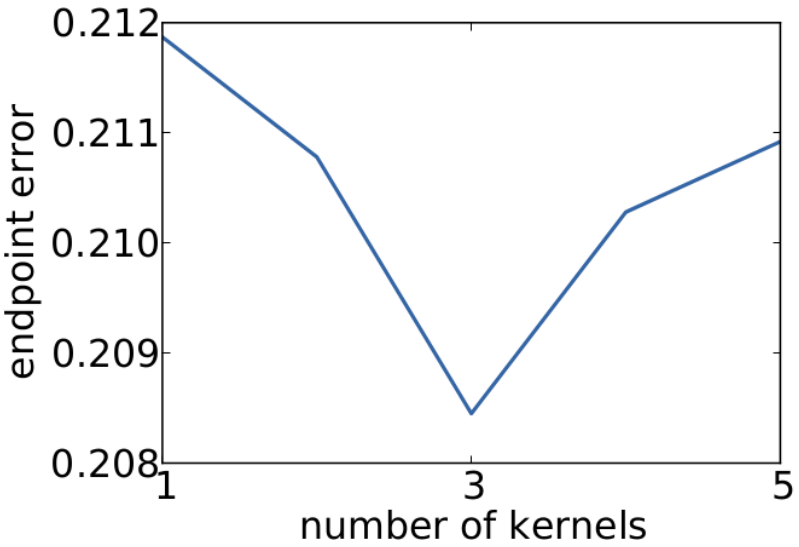
References

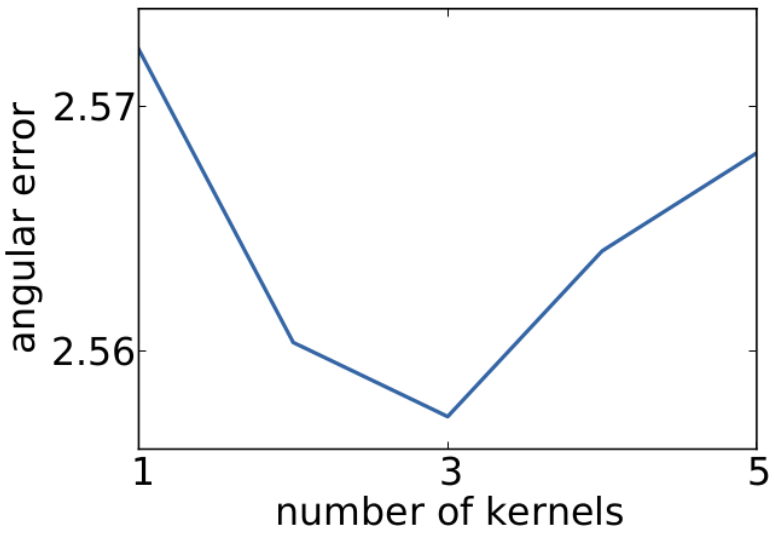
- Philipp Krähenbühl and Vladlen Koltun: Efficient Nonlocal Regularization for Optical Flow. European Conference on Computer Vision, 2012.
- N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert: Highly accurate optic flow computation with theoretically justified warping. International Journal of Computer Vision, 2006.
- A. Adams, J. Baek, and M. A. Davis: Fast high-dimensional filtering using the permutohedral lattice. Computer Graphics Forum, 2010.

Approximation of Charbonnier penalizer for $K = 3$

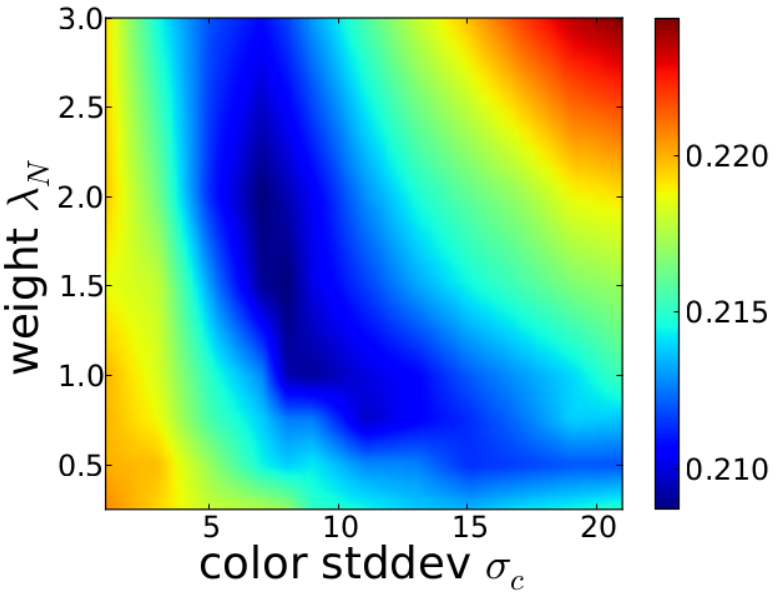


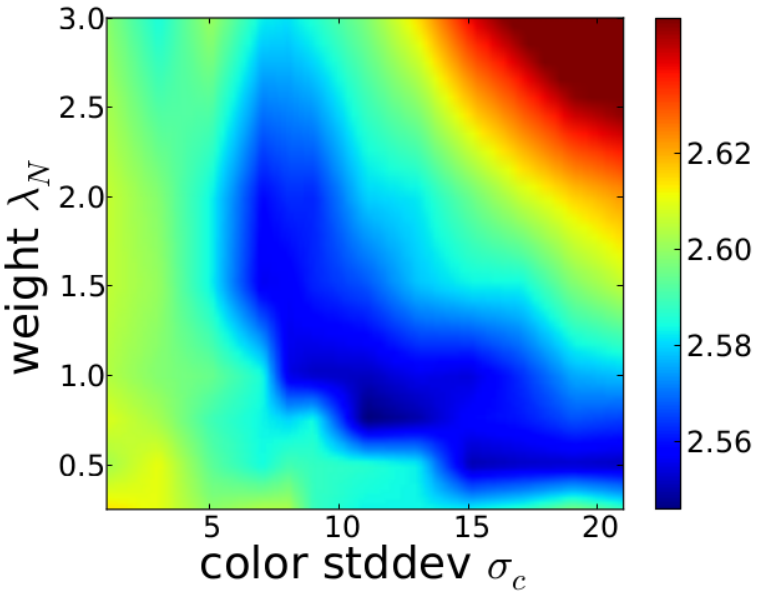
Average endpoint and angular error



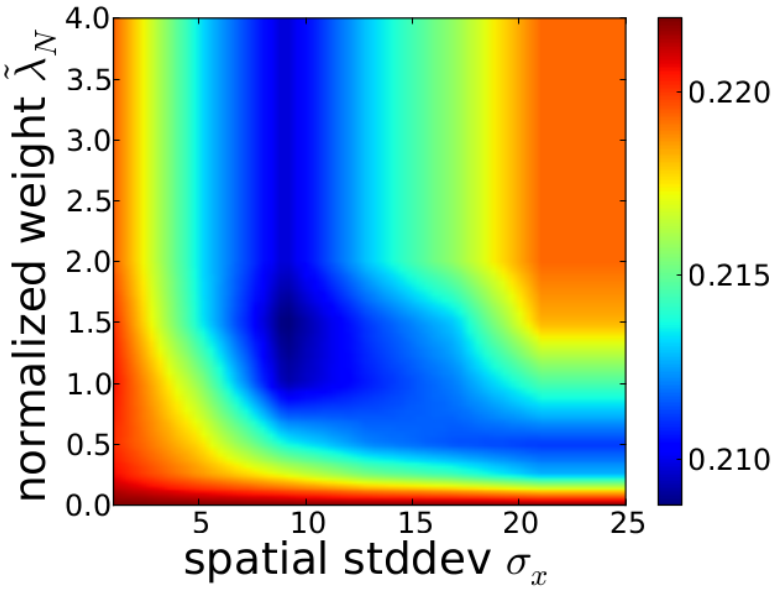


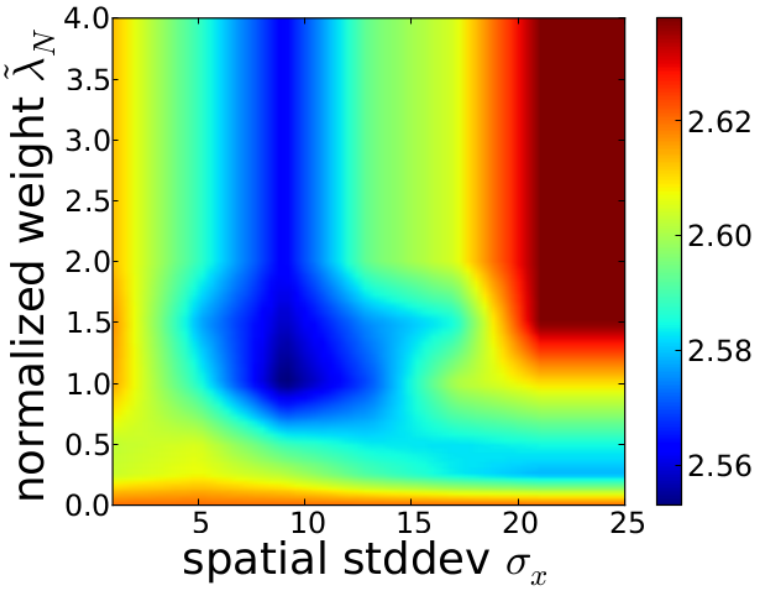
endpoint and angular errors for σ_c and λ_N





endpoint and angular errors for σ_x and λ_N

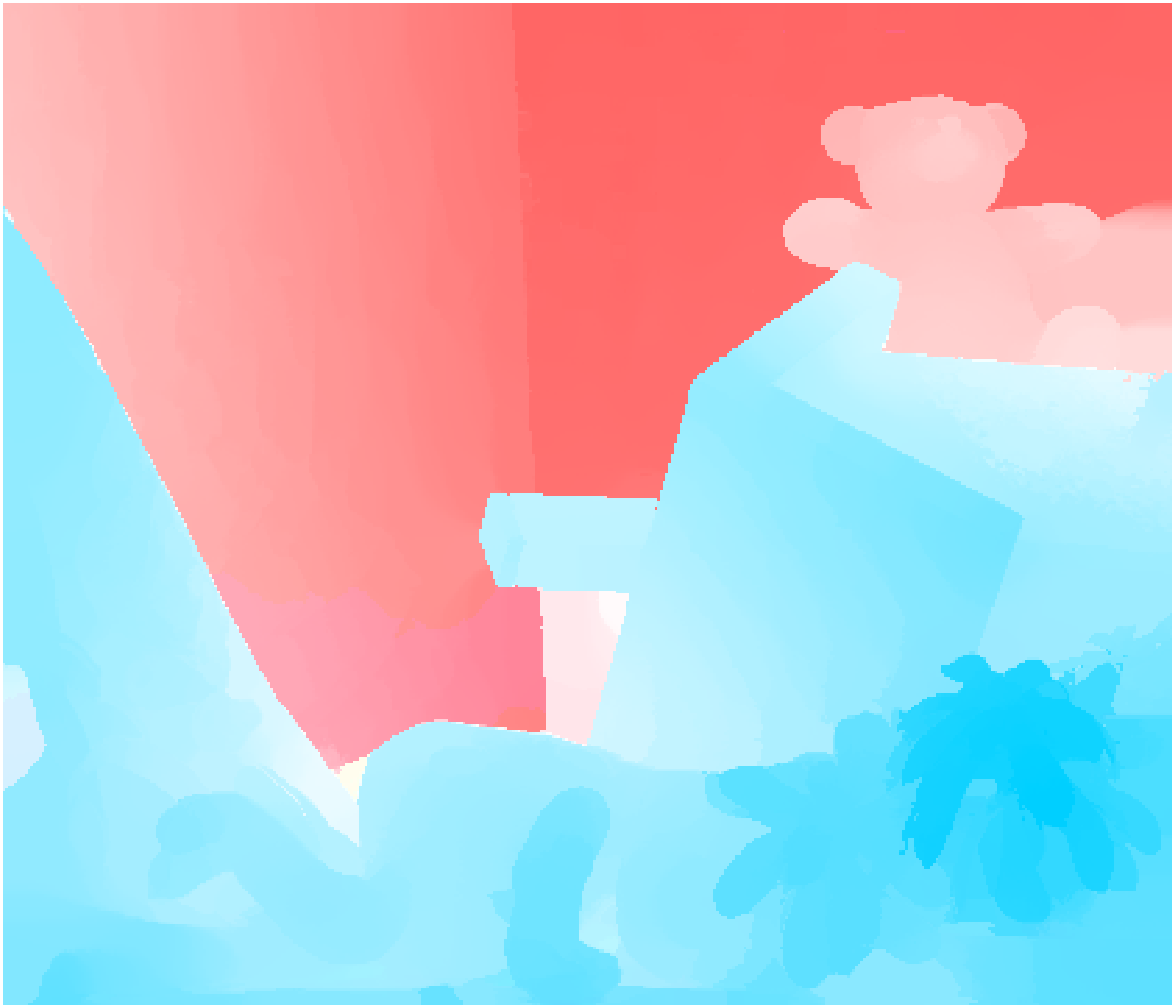




first image frame

ground truth flow

result without nonlocal regularisation



result with nonlocal regularisation