Guided Image Super-Resolution:

A New Technique for Photogeometric Super-Resolution in Hybrid 3-D Range Imaging

Florin C. Ghesu[1], Thomas Köhler[1,2], Sven Haase[1], Joachim Hornegger[1,2]
04.09.2014
[1] Pattern Recognition Lab
[2] Erlangen Graduate School in Advanced Optical Technologies (SAOT)

SAOT erlangen graduate school in advanced optical technologies

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Outline

Introduction

Proposed Guided Super-Resolution
- Bayesian Modeling
- Modeling the Image Formation Process
- Numerical Optimization

Experiments and Results
- Simulated Data
- Real Data

Summary and Conclusion

Introduction

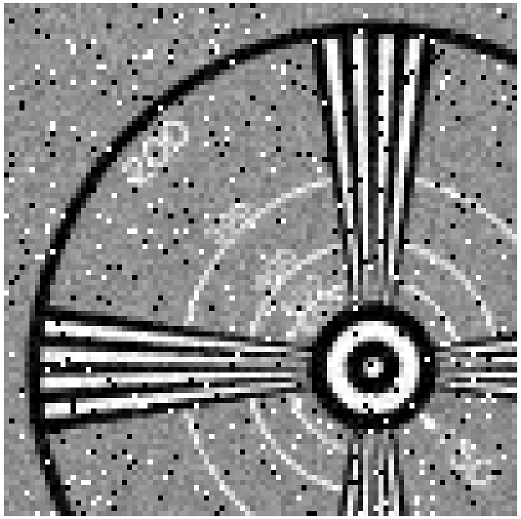
erlangen
graduate school in
advanced optical technologies

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

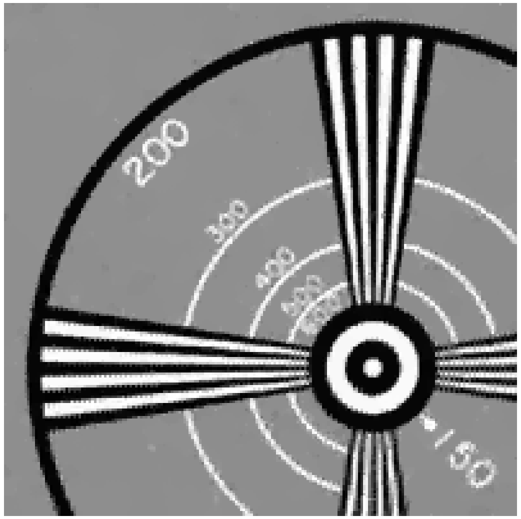
Single-Sensor Super-Resolution

- Reconstruct high-resolution image from multiple low-resolution frames
- Exploit subpixel motion present in low-resolution image sequence
- Conventional algorithms only applicable to single modality (sensor)

$\rightarrow$ **Single-sensor super-resolution**



Low-resolution



$2 \times$ Super-resolved

Multi-Sensor Super-Resolution

- **Main question:** Can super-resolution do a better job if we consider data from multiple sensors?

  $\rightarrow$ **Yes**, if we model dependencies/correlations between the sensors

- Applications:
  - **Hybrid range imaging:** 3-D range data augmented with photometric information
  - Multispectral imaging
  - Other hybrid imaging setups, e. g. PET/CT or PET/MR

  $\rightarrow$ **Multi-sensor super-resolution**

Related Work (in Hybrid Range Imaging)

- Single-sensor super-resolution applied to range images [1] [2]
  - Adopt techniques to range images originally introduced for color images
  - **Limitation**: does not exploit complementary photometric information
- Multi-sensor super-resolution for range images guided by photometric data
  - Guidance for motion estimation in presence of highly undersampled range data [3]
  - Adaptive regularization driven by color images [4]
  - **Limitation**: requires high-quality photometric information
- **Our contribution:**
  - New regularization technique to guide range super-resolution by photometric data
  - Super-resolved photometric data as by-product (photogeometric super-resolution)

[1] S. Schuon et al., (2008), *High-quality scanning using time-of-flight depth superresolution*, CVPR 2008
[2] S. Schuon et al., (2009), *LidarBoost: Depth superresolution for ToF 3D shape scanning*, CVPR 2009
[3] T. Köhler et al., (2013), *ToF Meets RGB: Novel Multi-Sensor Super-Resolution for Hybrid 3-D Endoscopy*, MICCAI 2013
[4] J. Park et al., (2010), *High quality depth map upsampling for 3D-TOF cameras*, ICCV 2011

Proposed Guided Super-Resolution

Bayesian Modeling of Multi-Sensor Super-Resolution

Single-sensor super-resolution:

- Given: sequence of low-resolution frames $\mathbf{y} = \begin{pmatrix} \mathbf{y}^{(1)} & \dots & \mathbf{y}^{(K)} \end{pmatrix}^\top$
- Consider data of one modality (range images)
- Maximum a posteriori (MAP) estimation to reconstruct the most probable high-resolution image **x**:

y, **x**

$$\begin{aligned} \hat{\mathbf{x}} &= \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) \\ &= \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \end{aligned} \tag{1}$$

Bayesian Modeling of Multi-Sensor Super-Resolution

Multi-sensor super-resolution with **independent channels**:

- Low-resolution range ($\mathbf{y}$) and photometric data ($\mathbf{p}$)
- High-resolution range ($\mathbf{x}$) and photometric data ($\mathbf{q}$)
- MAP estimation:

$$\begin{aligned}\hat{\mathbf{x}}, \hat{\mathbf{q}} &= \arg\max_{\mathbf{x},\mathbf{q}} p(\mathbf{x}, \mathbf{q} | \mathbf{y}, \mathbf{p}) \\ &= \arg\max_{\mathbf{x},\mathbf{q}} \underbrace{p(\mathbf{y}|\mathbf{x})p(\mathbf{p}|\mathbf{q})}_{\text{data likelihood}} \underbrace{p(\mathbf{x})p(\mathbf{q})}_{\text{prior}}\end{aligned} \tag{2}$$

$\rightarrow$ Single-sensor super-resolution applied to each channel

p, **q**

y, **x**

Bayesian Modeling of Multi-Sensor Super-Resolution

Multi-sensor super-resolution with **dependent channels**:

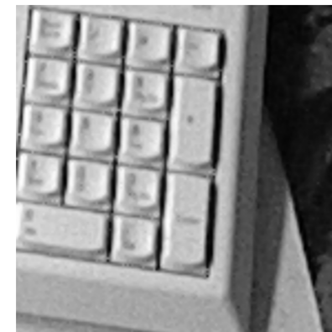
- The sensors „see“ the same scene
- Extend the MAP estimation:

$$\hat{\mathbf{x}}, \hat{\mathbf{q}} = \arg\max_{\mathbf{x},\mathbf{q}} p(\mathbf{x}, \mathbf{q} | \mathbf{y}, \mathbf{p}) = \arg\max_{\mathbf{x},\mathbf{q}} p(\mathbf{y}, \mathbf{p} | \mathbf{x}, \mathbf{q}) \, p(\mathbf{x}, \mathbf{q}) \tag{3}$$

Joint density for both modalities to model prior:

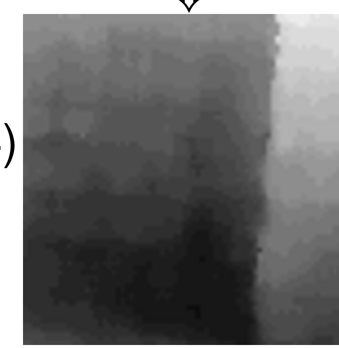
$$p(\mathbf{x}, \mathbf{q}) = \underbrace{p(\mathbf{x}) p(\mathbf{q} | \mathbf{x})}_{\text{dependencies}} \tag{4}$$

- How to model $p(\mathbf{y}, \mathbf{p} | \mathbf{x}, \mathbf{q})$, $p(\mathbf{x})$ and $p(\mathbf{q} | \mathbf{x})$?



p, q

$\Updownarrow$



y, x

Modeling the Image Formation Process

- Mathematical model $\mathcal{M}$ to describe formation of k-th low-resolution frame ($\mathbf{y}^{(k)}$ and $\mathbf{p}^{(k)}$) from high-resolution image ($\mathbf{x}$ and $\mathbf{q}$)

$$\mathcal{M}_x : \mathbf{x} \mapsto \mathbf{y}^{(k)} \quad \text{(range data)}$$
$$\mathcal{M}_q : \mathbf{q} \mapsto \mathbf{p}^{(k)} \quad \text{(photometric data)}$$

- Generative model for range and photometric data:

$$\begin{pmatrix} \mathbf{y}^{(k)} \\ \mathbf{p}^{(k)} \end{pmatrix} = \begin{pmatrix} \gamma_m^{(k)} \mathbf{W}_y^{(k)} & \mathbf{0} \\ \mathbf{0} & \eta_m^{(k)} \mathbf{W}_p^{(k)} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{q} \end{pmatrix} + \begin{pmatrix} \gamma_a^{(k)} \mathbf{1} \\ \eta_a^{(k)} \mathbf{1} \end{pmatrix} \tag{5}$$

  - $\mathbf{W}_y^{(k)}$ and $\mathbf{W}_p^{(k)}$ (system matrices) model subpixel motion, blur and subsampling
  - $\gamma_m^{(k)}$ and $\gamma_a^{(k)}$ models out-of-plane motion for range data
  - $\eta_m^{(k)}$ and $\eta_a^{(k)}$ models additive/multiplicative photometric differences

Joint Energy Minimization

- Log-likelihood function:

$$\begin{aligned}\hat{\mathbf{x}}, \hat{\mathbf{q}} &= \arg\min_{\mathbf{x},\mathbf{q}} \{-\log p(\mathbf{x}, \mathbf{q}|\mathbf{y}, \mathbf{p})\} \\ &= \arg\min_{\mathbf{x},\mathbf{q}} \{-\log (p(\mathbf{y}, \mathbf{p}|\mathbf{x}, \mathbf{q}) p(\mathbf{x}) p(\mathbf{q}|\mathbf{x}))\}\end{aligned} \tag{6}$$

- Formulation as unconstrained energy minimization problem:

$$(\hat{\mathbf{x}}, \hat{\mathbf{q}}) = \arg\min_{\mathbf{x},\mathbf{q}} \left\{ \underbrace{F_{\text{data}}(\mathbf{x}, \mathbf{q})}_{\text{Data likelihood: } p(\mathbf{y},\mathbf{p}|\mathbf{x},\mathbf{q})} + \underbrace{R_{\text{smooth}}(\mathbf{x}, \mathbf{q}) + R_{\text{correlate}}(\mathbf{x}, \mathbf{q})}_{\text{Prior: } p(\mathbf{x})p(\mathbf{p}|\mathbf{x})} \right\} \tag{7}$$

  We use photometric data to guide range data as modeled by $R_{\text{correlate}}(\mathbf{x}, \mathbf{q})$

- Joint optimization performed in cyclic coordinate descent scheme

Anatomy of the Objective Function

$$(\hat{\mathbf{x}}, \hat{\mathbf{q}}) = \arg\min_{\mathbf{x},\mathbf{q}} \left\{ F_{\text{data}}(\mathbf{x}, \mathbf{q}) + R_{\text{smooth}}(\mathbf{x}, \mathbf{q}) + R_{\text{correlate}}(\mathbf{x}, \mathbf{q}) \right\}$$

Data fidelity term for range and photometric data:

$$F_{\text{data}}(\mathbf{x}, \mathbf{q}) = \sum_{i=1}^{KN_y} \beta_{y,i} r_{y,i}(\mathbf{x})^2 + \sum_{i=1}^{KN_p} \beta_{p,i} r_{p,i}(\mathbf{q})^2, \tag{8}$$

- Residual error to measure data fidelity:

$$\begin{aligned} \mathbf{r}_y^{(k)} &= \mathbf{y}^{(k)} - \gamma_m^{(k)} \mathbf{W}_y^{(k)} \mathbf{x} - \gamma_a^{(k)} \mathbf{1} \\ \mathbf{r}_p^{(k)} &= \mathbf{p}^{(k)} - \eta_m^{(k)} \mathbf{W}_p^{(k)} \mathbf{q} - \eta_a^{(k)} \mathbf{1}. \end{aligned} \tag{9}$$

- $\beta_{y,i}$ and $\beta_{p,i}$ are confidence maps (estimated in our optimization algorithm)

Anatomy of the Objective Function

$$(\hat{\mathbf{x}}, \hat{\mathbf{q}}) = \arg\min_{\mathbf{x},\mathbf{q}} \left\{ F_{\text{data}}(\mathbf{x}, \mathbf{q}) + R_{\text{smooth}}(\mathbf{x}, \mathbf{q}) + R_{\text{correlate}}(\mathbf{x}, \mathbf{q}) \right\}$$

Smoothness regularization for range and photometric data:

$$R_{\text{smooth}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \lambda_x R(\mathbf{x}) + \lambda_q R(\mathbf{q}) \tag{10}$$

- Edge preserving regularization weighted by $\lambda_x \geq 0$ and $\lambda_q \geq 0$
- We use the bilateral total variation (BTV) [5]:

$$R(\mathbf{z}) = \sum_{i=-P}^{P} \sum_{j=-P}^{P} \alpha^{|i|+|j|} \left\| \mathbf{z} - \mathbf{S}_v^i \mathbf{S}_h^j \mathbf{z} \right\|_1 \tag{11}$$

Calculates BTV for local neighborhood of radius P with weighting factor α where $\mathbf{S}_v^i$ and $\mathbf{S}_h^j$ models vertical/horizontal shifts of $\mathbf{z}$

[5] S. Farsiu et al., *Fast and Robust Multi-Frame Super-Resolution*, IEEE TIP, 2004

Anatomy of the Objective Function

$$(\hat{\mathbf{x}}, \hat{\mathbf{q}}) = \arg\min_{\mathbf{x},\mathbf{q}} \left\{ F_{\text{data}}(\mathbf{x}, \mathbf{q}) + R_{\text{smooth}}(\mathbf{x}, \mathbf{q}) + R_{\text{correlate}}(\mathbf{x}, \mathbf{q}) \right\}$$

Interdependence regularization between both modalities:

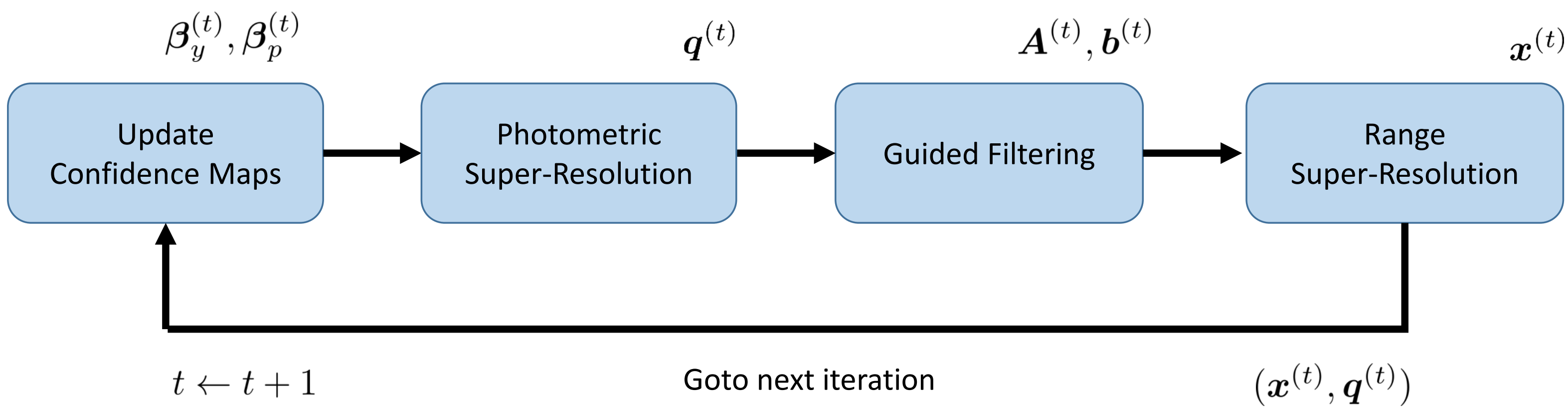
$$R_{\text{correlate}}(\mathbf{x}, \mathbf{q}) = \lambda_c ||\mathbf{x} - \mathbf{A}\mathbf{q} - \mathbf{b}||_2^2 \tag{12}$$

Local (patch-wise) linear correlation model defined by **guided filtering** [6]

- **A** (diagonal matrix) and **b** are guided filter coefficients (estimated in our optimization algorithm)
- $\lambda_c \geq 0$ indicates how strong photometric data guides the range data

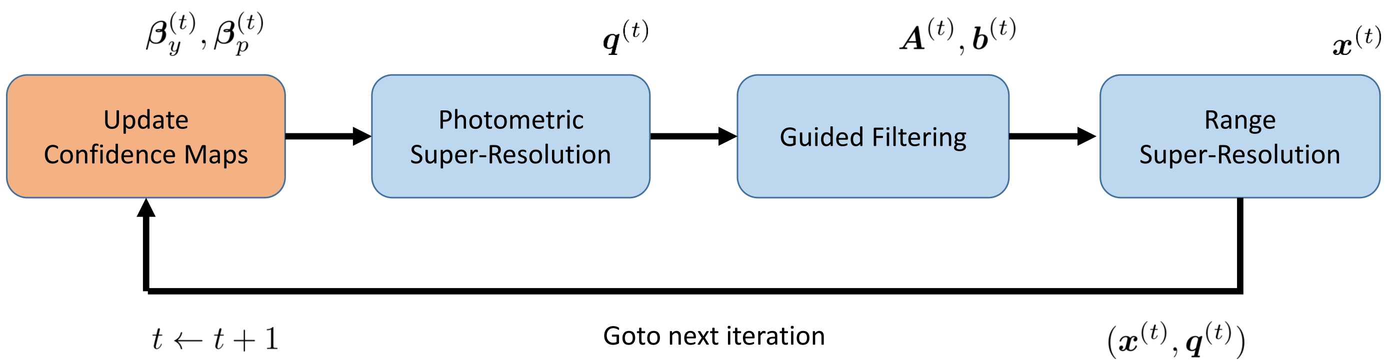
[6] K. He et al., *Guided Image Filtering*, IEEE PAMI, 2013

Numerical Optimization



- We employ iteratively re-weighted least squares (IRLS) optimization to reconstruct super-resolved range and photometric data
  $\rightarrow$ Iteration sequence: let $(\mathbf{x}^{(t)}, \mathbf{q}^{(t)})$ be the estimates at iteration t
- Guided filter coefficients (interdependence regularization) and confidence maps (data fidelity term) are iteratively updated

Numerical Optimization

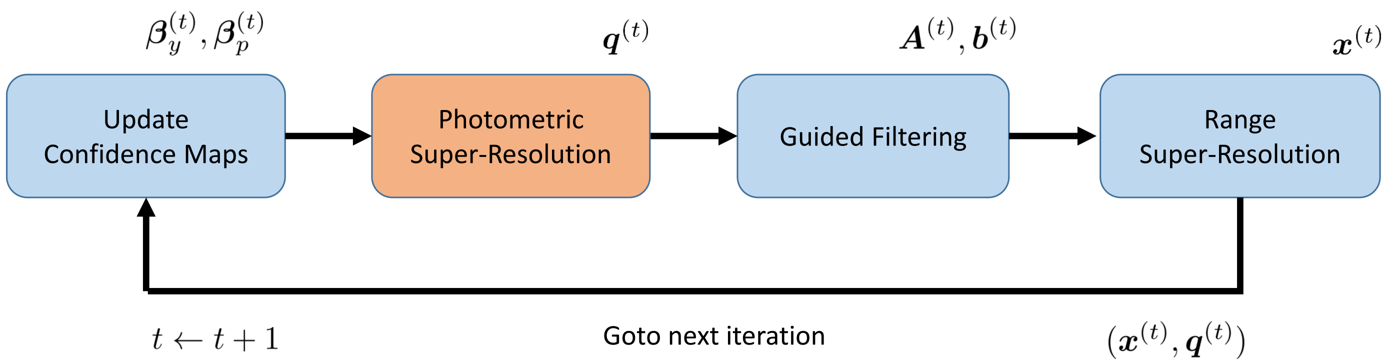


Step 1 (confidence maps): derive from the residual error for $(\mathbf{x}^{(t)}, \mathbf{q}^{(t)})$

$$\beta_{y,i}^{(t)} = \begin{cases} 1 & \text{if } |\mathbf{r}_{y,i}^{(t)}| \leq \epsilon_y \\ \frac{\epsilon_y}{|\mathbf{r}_{y,i}^{(t)}|} & \text{otherwise} \end{cases} \qquad \beta_{p,i}^{(t)} = \begin{cases} 1 & \text{if } |\mathbf{r}_{p,i}^{(t)}| \leq \epsilon_p \\ \frac{\epsilon_p}{|\mathbf{r}_{p,i}^{(t)}|} & \text{otherwise} \end{cases} \tag{13}$$

- Assign smaller confidence to observation with higher residual
- ϵ_y and ϵ_p are estimated from the median absolute deviation of the residual

Numerical Optimization

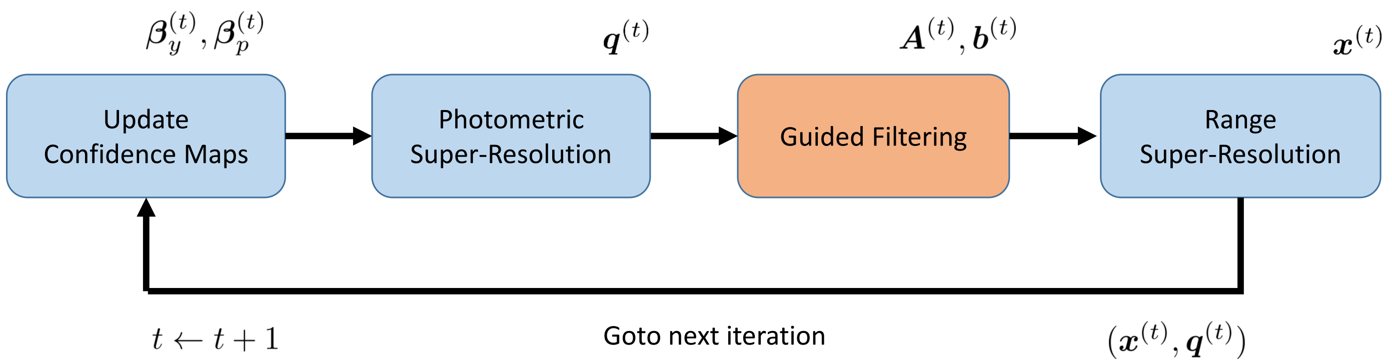


Step 2 (photometric super-resolution): update $\mathbf{q}^{(t-1)}$ to $\mathbf{q}^{(t)}$ for fixed $\mathbf{x}$

$$\mathbf{q}^{(t)} = \arg\min_{\mathbf{q}} \left\{ F_{\text{data}}(\mathbf{x}, \mathbf{q}) + R_{\text{smooth}}(\mathbf{x}, \mathbf{q}) \right\}_{\mathbf{x}=\mathbf{x}^{(t-1)}} \tag{14}$$

- Interdependence regularization not used (photometric data guides range data but not vice versa)
- Convex optimization problem solved by scaled conjugate gradient method

Numerical Optimization



Step 3 (guided filtering): compute filter coefficients $\mathbf{A}^{(t)}$ and $\mathbf{b}^{(t)}$ [7]

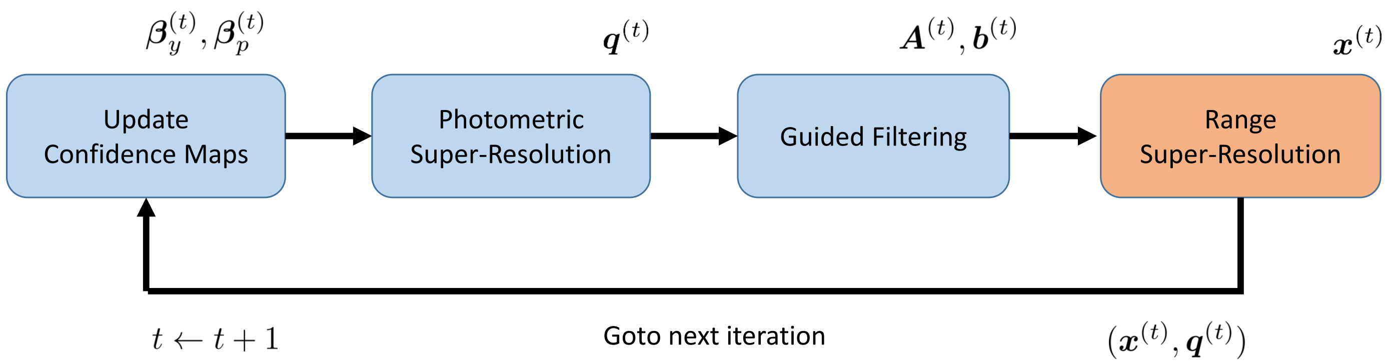
$$\tilde{A}_{k,k} = \frac{\frac{1}{|\omega_k|}\sum_{i \in \omega_k} q_i x_i - \mathrm{E}_{\omega_k}(\mathbf{q})\mathrm{E}_{\omega_k}(\mathbf{x})}{\mathrm{Var}_{\omega_k}(\mathbf{q}) + \epsilon} \quad (15)$$

$$\tilde{b}_k = \mathrm{E}_{\omega_k}(\mathbf{x}) - \tilde{A}_{k,k}\mathrm{E}_{\omega_k}(\mathbf{q}) \quad (16)$$

Filter parameters: $|\omega_k|$ (kernel size) and ϵ (regularization parameter)

[7] K. He et al., *Guided Image Filtering*, IEEE PAMI, 2013

Numerical Optimization



Step 4 (range super-resolution): update $\mathbf{x}^{(t-1)}$ to $\mathbf{x}^{(t)}$ for fixed $\mathbf{q}$

$$\mathbf{x}^{(t)} = \arg\min_{\mathbf{x}} \left\{ F_{\text{data}}(\mathbf{x}, \mathbf{q}) + R_{\text{smooth}}(\mathbf{x}, \mathbf{q}) + R_{\text{correlate}}(\mathbf{x}, \mathbf{q}) \right\}_{\mathbf{q}=\mathbf{q}^{(t)}} \tag{17}$$

- Use filter coefficients $\mathbf{A}^{(t)}$ and $\mathbf{b}^{(t)}$ for interdependence regularization
- Convex optimization problem solved by scaled conjugate gradient method

Experiments and Results

erlangen graduate school in advanced optical technologies

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Experiments and Results

Experiments:

- Simulated data with known ground truth
- Real datasets (Microsoft's Kinect)

Compared methods:

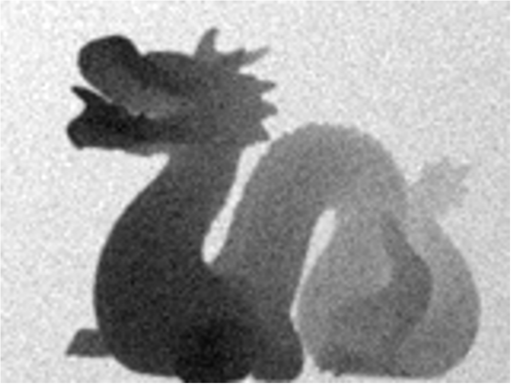
- MAP super-resolution with L_2 norm model (applied to each modality separately)
- Robust super-resolution based on L_1 norm model (applied to each modality separately) [8]
- Proposed method for photogeometric super-resolution

Motion estimation: optical flow on photometric data (employed for all super-resolution methods) [9]

[8] S. Schuon et al., (2008), *High-quality scanning using time-of-flight depth superresolution*, CVPR 2008

[9] T. Köhler et al., (2013), *ToF Meets RGB: Novel Multi-Sensor Super-Resolution for Hybrid 3-D Endoscopy*, MICCAI 2013

Simulated Data: Experimental Setup



Low-resolution range

Ground truth

- Ground truth range/photometric data simulated with 640×480 px [10]
- Simulation of Gaussian PSF and subsampling (factor: 4) to generate low-resolution data
- 4 datasets: 10 image sequences ($K = 31$ low-resolution frames per sequence)

[10] using the range imaging toolkit (RITK): http://www5.cs.fau.de/research/software/range-imaging-toolkit-ritk/

Simulated Data: Results

Evaluation for range data:

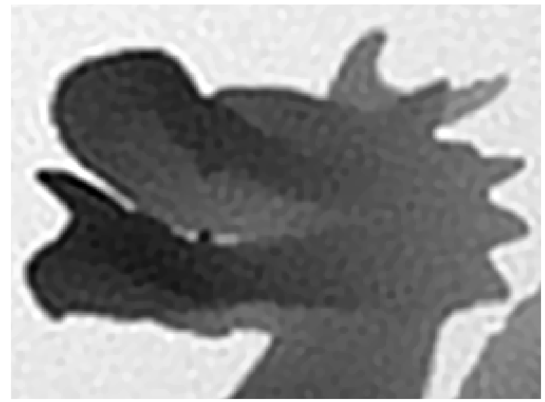
MAP (L_2 norm)

MAP (L_1 norm)

Proposed

Simulated Data: Results

MAP (L_2 norm)



MAP (L_1 norm)

Proposed

Ground truth

Simulated Data: Results

Evaluation for photometric data:

Low-resolution

Guided approach

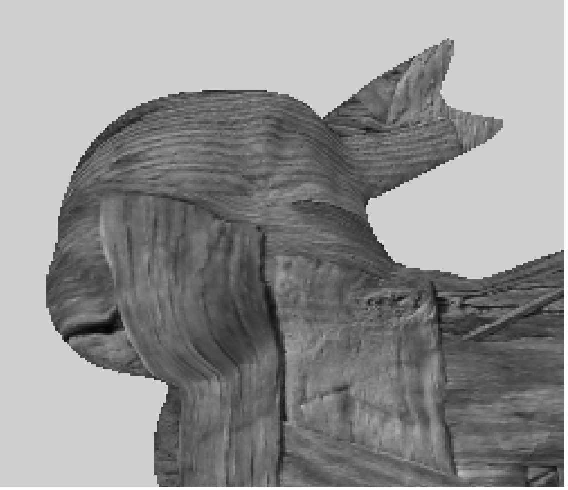
Ground truth

Simulated Data: Results

Evaluation for photometric data:

Low-resolution

Guided approach



Ground truth

Simulated Data: Results

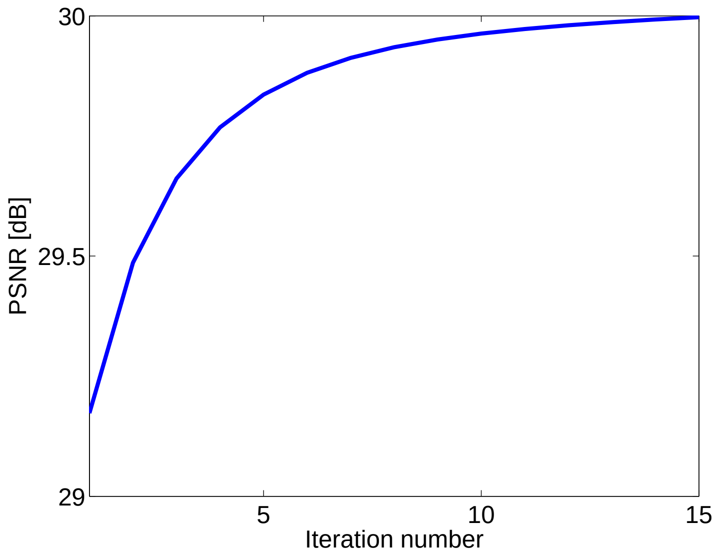
- Evaluation of peak-signal-to-noise ratio (PSNR) and structural similarity (SSIM) for **range and photometric data**:
- All results averaged over $n = 10$ test sequences

	Sequence	**Low-resolution**	**MAP - L_1**	**MAP - L_2**	**Proposed**
Range	Bunny-1	32.78 (0.96)	34.10 (0.96)	34.05 (0.97)	**35.01 (0.98)**
	Bunny-2	31.29 (0.94)	32.84 (0.95)	33.22 (0.97)	**33.34 (0.98)**
	Dragon-1	24.63 (0.57)	27.68 (0.72)	28.71 (0.84)	**30.00 (0.91)**
	Dragon-2	27.14 (0.75)	29.09 (0.84)	29.76 (0.93)	**30.80 (0.95)**
Photom.	Bunny-1	28.48 (0.79)	29.82 (0.87)	29.79 (0.87)	29.79 (0.88)
	Bunny-2	30.05 (0.81)	31.35 (0.86)	31.42 (0.86)	31.43 (0.86)
	Dragon-1	23.34 (0.65)	24.25 (0.72)	24.24 (0.71)	24.27 (0.72)
	Dragon-2	24.65 (0.66)	25.60 (0.72)	25.51 (0.70)	25.60 (0.72)

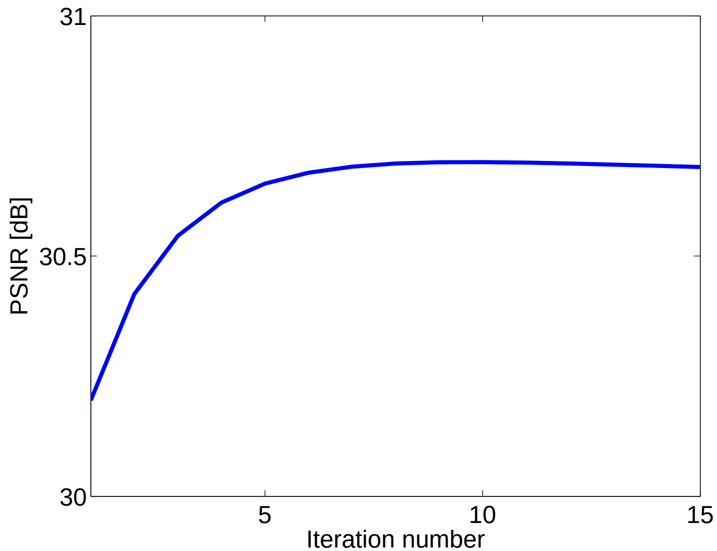
$\rightarrow$ Improved PSNR/SSIM for range data and competitive results for photometric data (photometric data guides range data, not vice versa)

Simulated Data: Results

Experimental analysis of the convergence: PSNR vs. IRLS iteration number



Dragon-1 dataset



Dragon-2 dataset

- Initialization: MAP super-resolution based on L_2 norm
- Typically converged after $\approx$ 10 IRLS iterations

Real Data: Experimental Setup

Microsoft's Kinect:

Photometric data

Range data

- Acquisition of real data using Microsoft's Kinect (640 $\times$ 480 px, 30 fps)
- Subpixel motion due to small shaking of the device
- Datasets with sequences of $K = 31$ frames (magnification factor: 4)

Real Data: Results

Evaluation for range data:

Low-resolution

MAP (L_2 norm)

MAP (L_1 norm)

Proposed

Real Data: Results

Evaluation for photometric data:

Low-resolution

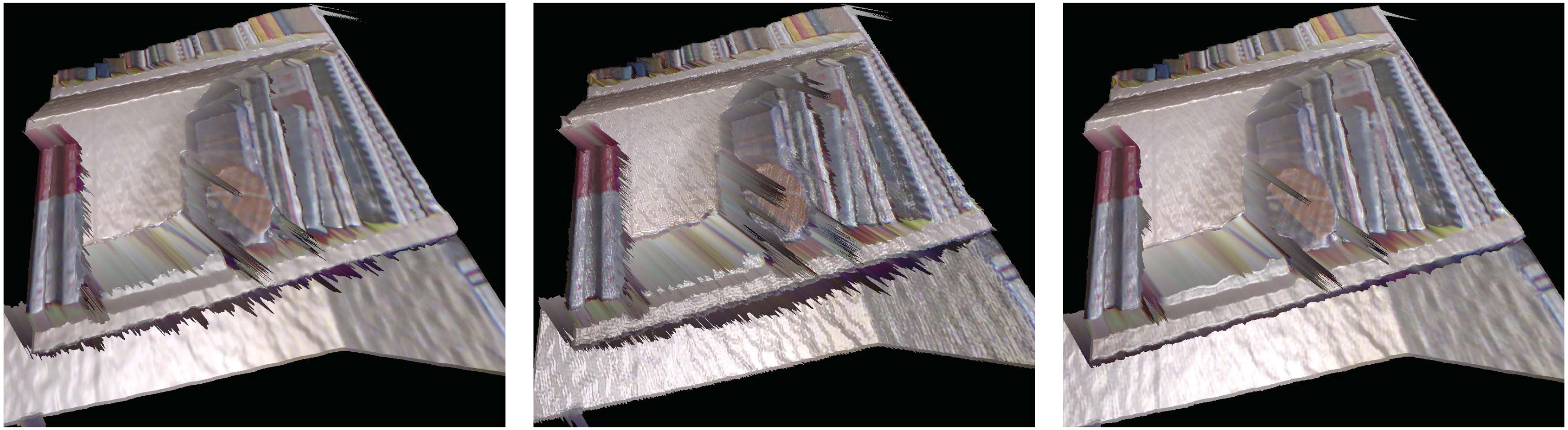
MAP (L_2 norm)

MAP (L_1 norm)

Proposed

Real Data: Results

Super-resolved 3-D mesh with color overlay:



MAP (L_2 norm) MAP (L_1 norm) Proposed

- Invalid pixels (due to occlusions) reconstructed by all methods
- Improved reconstruction of edges by our method

Summary and Conclusion

SAOT erlangen graduate school in advanced optical technologies

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Summary and Conclusion

- **Novel interdependence regularization** to guide range super-resolution by photometric data
- **Photogeometric resolution enhancement:** super-resolve range and photometric data in a joint framework
- **Robust image reconstruction** based on IRLS optimization

Outlook: Adaption/generalization for other sensors and hybrid imaging setups, e. g.

- Time-of-Flight imaging (range + amplitude data)
- RGB-D imaging to handle multiple color channels
- Multispectral imaging

Supplementary Material

A super-resolution toolbox (Matlab & MEX/C++) and datasets used for our experiments are available on our webpage:

http://www5.cs.fau.de/research/data

Thank you very much for your attention!

Acknowledgments

Thank you very much for the support of this work

Deutsche Forschungsgemeinschaft

erlangen graduate school in advanced optical technologies