Learning Joint Demosaicing and Denoising Based on Sequential Energy Minimization

Teresa Klatzer[1], Kerstin Hammernik[1], Patrick Knöbelreiter[1], Thomas Pock[1,2]

[1] Institute of Computer Graphics and Vision
Graz University of Technology

[2] Digital Safety and Security Department
AIT Austrian Institute of Technology

Der Wissenschaftsfonds.

ICCP 2016, Evanston IL

Image Aquisition

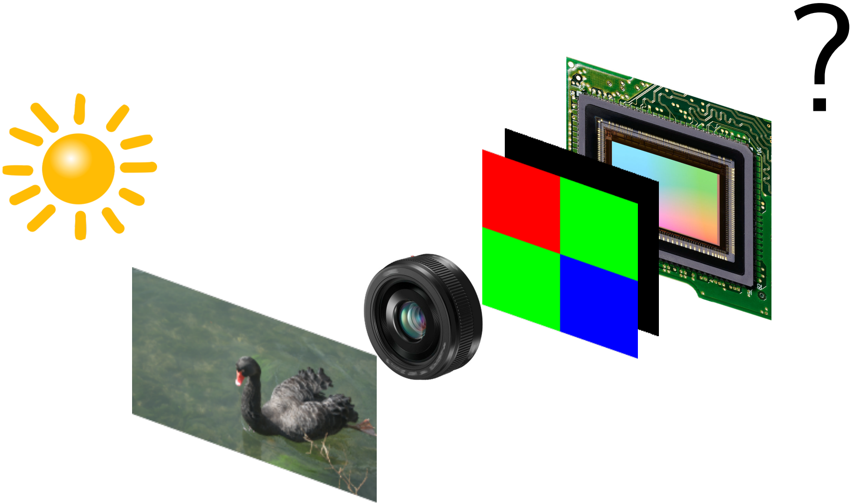
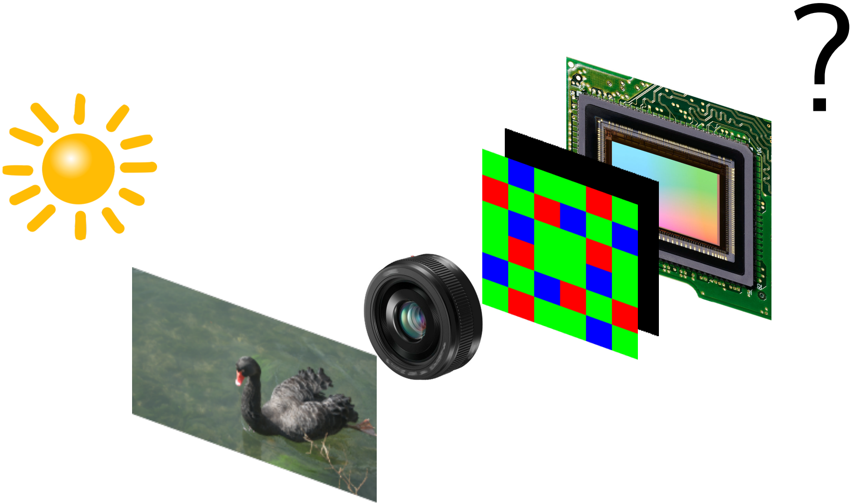
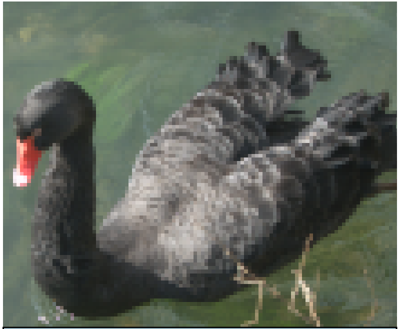


Image Aquisition



Imaging Pipeline



Input scene

Image Acquisition

RAW mosaiced image

Imaging Pipeline

RAW mosaiced image

Mosaic Operator A

Mosaic image
linear RGB space

Imaging Pipeline

Mosaic image
linear RGB space

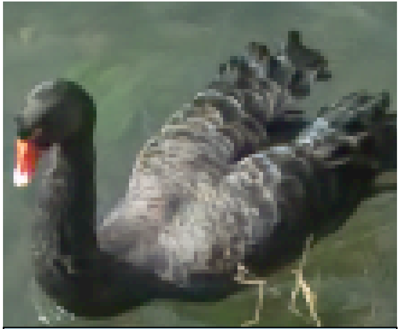
Demosaicing and Denoising

Demosaiced image
linear RGB space

Imaging Pipeline

Demosaiced image
linear RGB space

Color and Gamma Correction



Fully developed image
sRGB space

Joint Demosaicing and Denoising

In summary,

Joint Demosaicing and Denoising

In summary,

- Demosaicing is the process of filling missing pixels to produce full color images

Joint Demosaicing and Denoising

In summary,

- Demosaicing is the process of filling missing pixels to produce full color images
- Depends on Color Filter Array (CFA) pattern

Joint Demosaicing and Denoising

In summary,

- Demosaicing is the process of filling missing pixels to produce full color images
- Depends on Color Filter Array (CFA) pattern
- Noise is always present (sensor noise)

Joint Demosaicing and Denoising

In summary,

- Demosaicing is the process of filling missing pixels to produce full color images
- Depends on Color Filter Array (CFA) pattern
- Noise is always present (sensor noise)

How to solve these problems all in one?

How this problem has been solved

- Older works: interpolation based, post-processing to get rid of artifacts, exploit edge information and color channel correlations
  - Spatially adaptive interpolation (LPA) [Paliy et al., 2007]
  - Contour stencils for edge-adaptive interpolation (CS) [Getreuer, 2009]
  - And many more!

How this problem has been solved

- Older works: interpolation based, post-processing to get rid of artifacts, exploit edge information and color channel correlations
  - Spatially adaptive interpolation (LPA) [Paliy et al., 2007]
  - Contour stencils for edge-adaptive interpolation (CS) [Getreuer, 2009]
  - And many more!
- Recent Works: Handle denoising and demosaicing jointly
  - Color correlations and non-local means (JMCDM) [Chang et al., 2015]
  - Reconstruction based on several image priors like TV, BM3D (FlexISP) [Heide et al., 2014]
  - Learning: Using Regression Tree Fields (RTF) [Khashabi et al., 2014]

How this problem has been solved

- Older works: interpolation based, post-processing to get rid of artifacts, exploit edge information and color channel correlations
  - Spatially adaptive interpolation (LPA) [Paliy et al., 2007]
  - Contour stencils for edge-adaptive interpolation (CS) [Getreuer, 2009]
  - And many more!
- Recent Works: Handle denoising and demosaicing jointly
  - Color correlations and non-local means (JMCDM) [Chang et al., 2015]
  - Reconstruction based on several image priors like TV, BM3D (FlexISP) [Heide et al., 2014]
  - Learning: Using Regression Tree Fields (RTF) [Khashabi et al., 2014]
    - Microsoft Demosaicing Data Base

Simple Model Assumptions May Fail

Main problems:

- Edge blur
- Zippering
- False color
- Aliasing
- How to deal with noise?

Simple Model Assumptions May Fail

Main problems:

- Edge blur
- Zippering
- False color
- Aliasing
- How to deal with noise?

Simple Model Assumptions May Fail

Main problems:

- Edge blur
- Zippering
- False color
- Aliasing
- How to deal with noise?

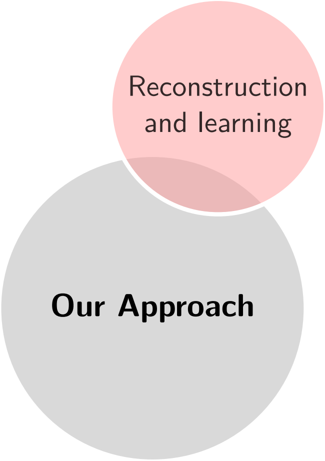
Simple Model Assumptions May Fail

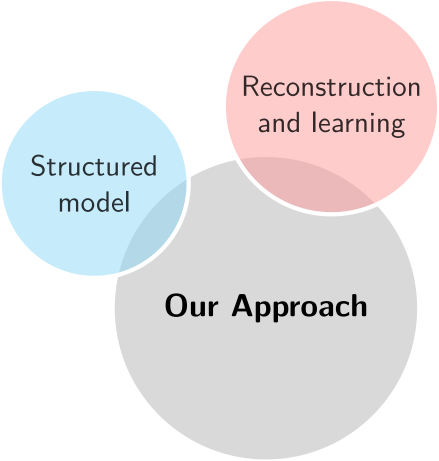
Main problems:

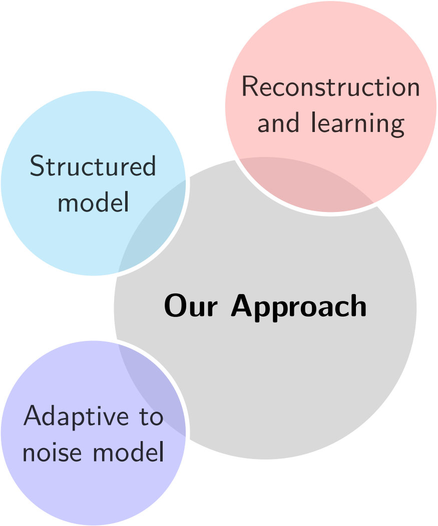
- Edge blur
- Zippering
- False color
- Aliasing
- How to deal with noise?

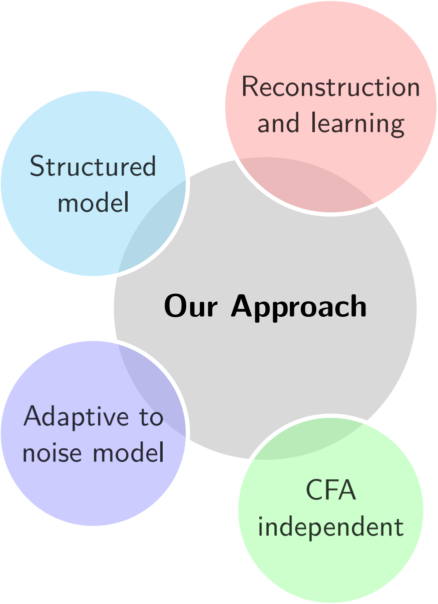
And even learning approaches have problems...

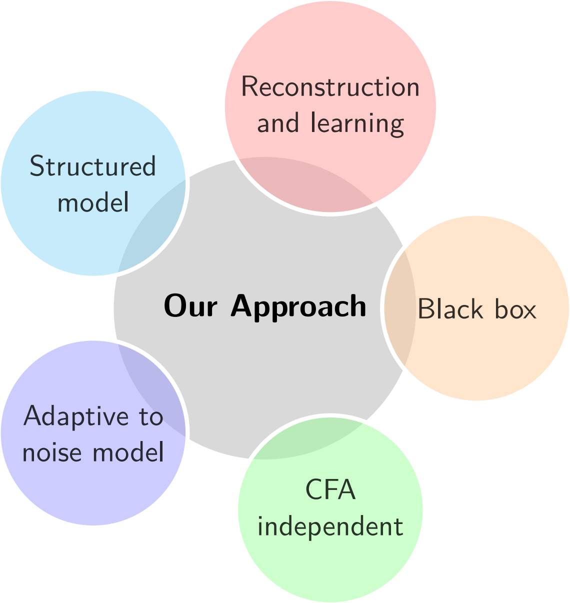
Our Approach


Reconstruction and learning
Our Approach


Reconstruction and learning
Structured model
Our Approach


Reconstruction and learning
Structured model
Our Approach
Adaptive to noise model


Reconstruction and learning
Structured model
Our Approach
Adaptive to noise model
CFA independent


Reconstruction and learning
Structured model
Our Approach
Black box
Adaptive to noise model
CFA independent

Variational Image Reconstruction

- In a **variational reconstruction** approach, we seek to find image u given the observation m such that

$$u \in \arg\min_{u} \{\mathcal{R}(u) + \lambda \mathcal{D}(u, m)\}$$

Variational Image Reconstruction

- In a **variational reconstruction** approach, we seek to find image u given the observation m such that

$$u \in \arg\min_u \left\{ \mathcal{R}(u) + \lambda \frac{1}{2} \| Au - m \|_2^2 \right\}$$

- Typically, these terms are fixed and handcrafted, with manually tuned parameters.

Variational Image Reconstruction

- In a **variational reconstruction** approach, we seek to find image u given the observation m such that

$$u \in \arg\min_u \left\{ \mathcal{R}(u) + \lambda \frac{1}{2} \|Au - m\|_2^2 \right\}$$

- Typically, these terms are fixed and handcrafted, with manually tuned parameters.
- In **our model**, we assume parametrized functions to serve as regularizer

$$u \in \arg\min_u \left\{ \sum_{i=1}^{N} \sum_{p=1}^{HW} \rho_i((k_i * u)_p) + \lambda \frac{1}{2} \|A\,u - m\|_2^2 \right\}$$

Variational Image Reconstruction

- In a **variational reconstruction** approach, we seek to find image u given the observation m such that

$$u \in \arg\min_u \left\{ \mathcal{R}(u) + \lambda \frac{1}{2} \|Au - m\|_2^2 \right\}$$

- Typically, these terms are fixed and handcrafted, with manually tuned parameters.
- In **our model**, we assume parametrized functions to serve as regularizer

$$u \in \arg\min_u \left\{ \sum_{i=1}^{N} \sum_{p=1}^{HW} \rho_i((k_i * u)_p) + \lambda \frac{1}{2} \|A\,u - m\|_2^2 \right\}$$

- and in a second extension, freely tunable functions for the data term

$$u \in \arg\min_u \left\{ \sum_{i=1}^{N} \sum_{p=1}^{HW} \rho_i((k_i * u)_p) + \lambda \sum_{p=1}^{HW} \Psi(A\,u - m)_p \right\}$$

Combining Successful Ideas

- Gray image restoration model: Nonlinear Reaction-Diffusion [Chen et al., 2015]
- Extend the model with a trainable color image prior (SEM)
- Extend the model with an adaptive data term (SEM+D)
- We have now a variational model of the form:

$$u \in \arg\min_{u} \left\{ \sum_{i=1}^{N} \sum_{p=1}^{HW} \rho_i \left(\sum_{c \in \{r,g,b\}} (k_{c,i} * u)_p \right) + \lambda \sum_{p=1}^{HW} \Psi(A\,u - m)_p \right\}$$

- Using the above equation we formalize a sequence of S gradient steps

$$u_c^s = u_c^{s-1} - \sum_{i=1}^{N} \bar{k}_{c,i}^s * \phi_i^s \left(\sum_{c \in \{r,g,b\}} (k_{c,i}^s * u^{s-1}) \right) - \lambda^s A^T \psi(A\,u^{s-1} - m)$$

Combining Successful Ideas

- Gray image restoration model: Nonlinear Reaction-Diffusion [Chen et al., 2015]
- Extend the model with a trainable color image prior (SEM)
- Extend the model with an adaptive data term (SEM+D)
- We have now a variational model of the form:

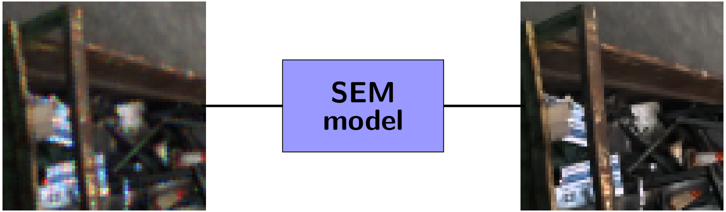
$$u \in \arg\min_u \left\{ \sum_{i=1}^{N} \sum_{p=1}^{HW} \rho_i \left(\sum_{c \in \{r,g,b\}} (k_{c,i} * u)_p \right) + \lambda \sum_{p=1}^{HW} \Psi(A\,u - m)_p \right\}$$

- Using the above equation we formalize a sequence of S gradient steps

$$u_c^s = u_c^{s-1} - \sum_{i=1}^{N} \bar{k}_{c,i}^s * \phi_i^s \left(\sum_{c \in \{r,g,b\}} (k_{c,i}^s * u^{s-1}) \right) - \lambda^s A^T \psi(A\,u^{s-1} - m)$$

Sequential Energy Minimization Model (SEM)

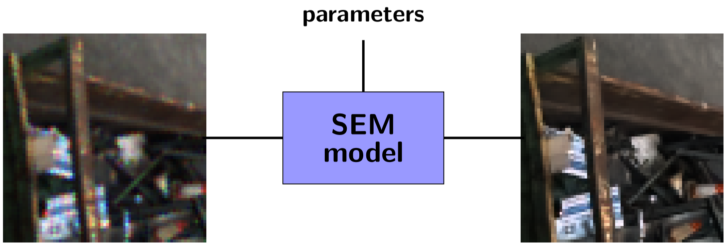
Parameter Learning



initialization
bilinear interpolation

reconstruction

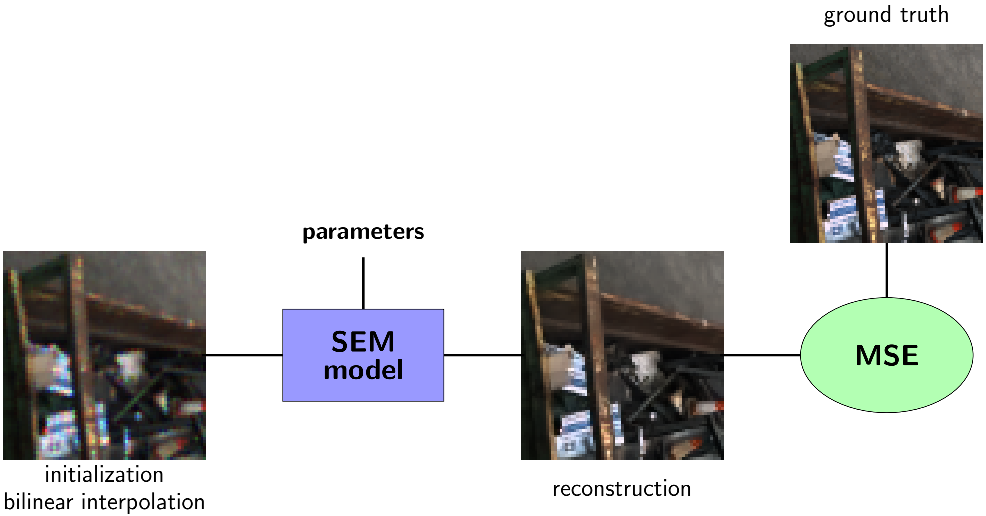
Parameter Learning



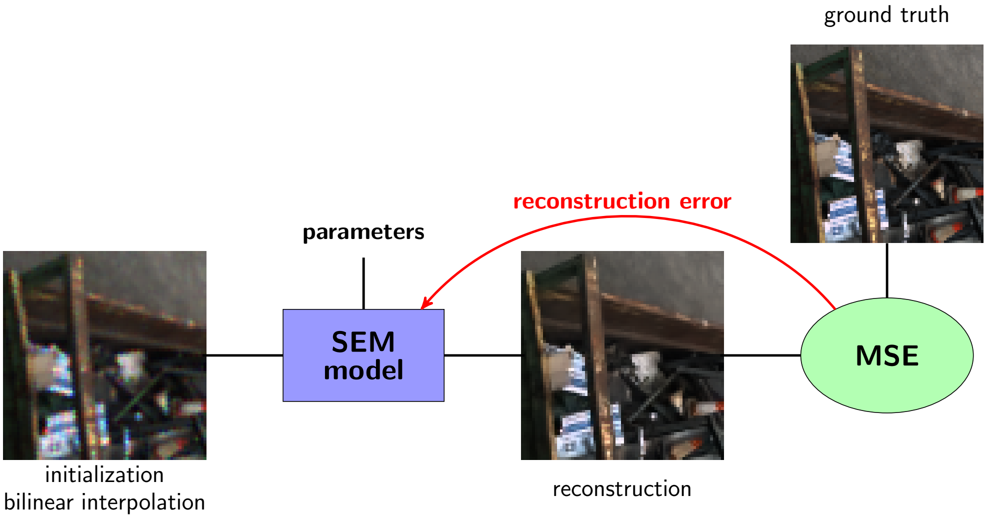
initialization
bilinear interpolation

reconstruction

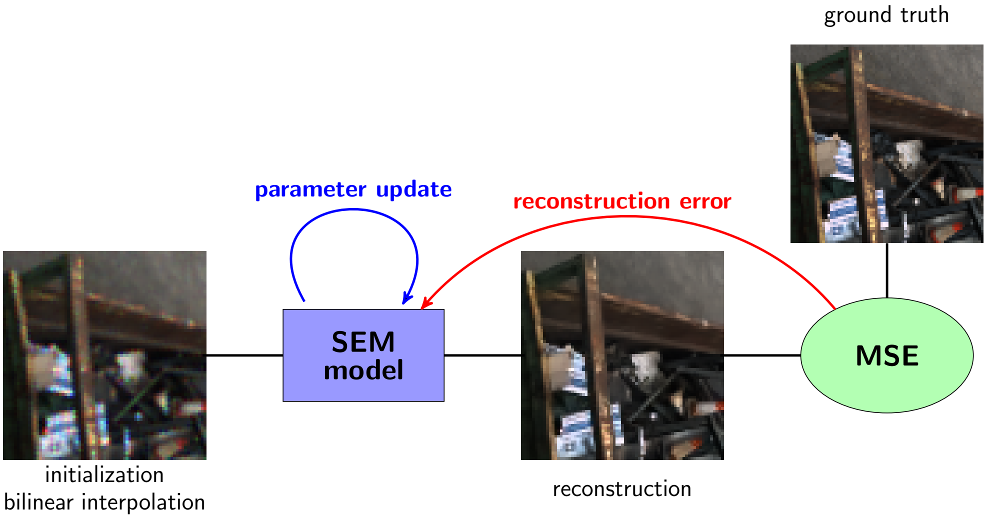
Parameter Learning



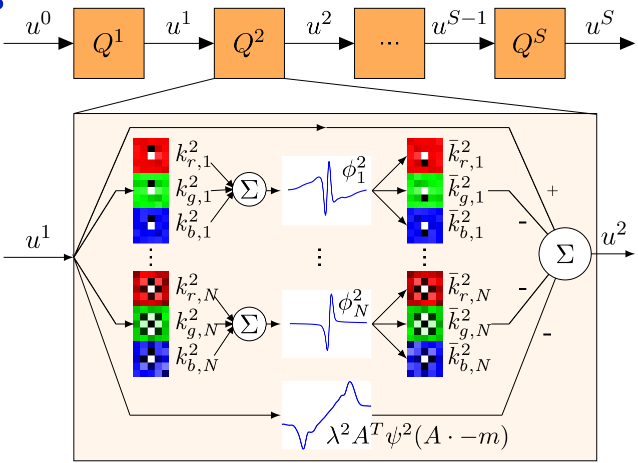
Parameter Learning



Parameter Learning

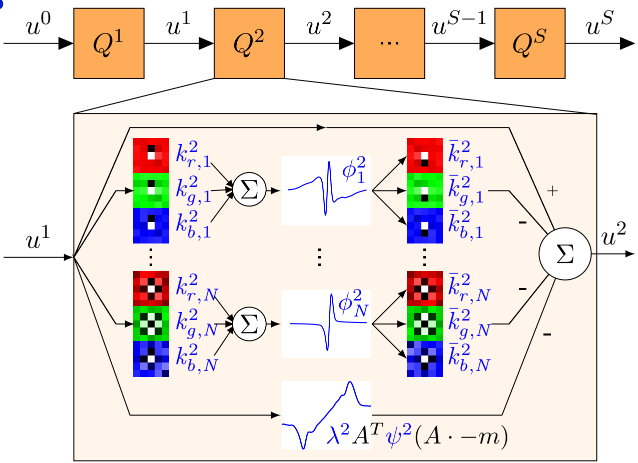


Gradient Steps



$$u_c^s = u_c^{s-1} - \sum_{i=1}^{N} \bar{k}_{c,i}^s * \phi_i^s \left(\sum_{c \in \{r,g,b\}} (k_{c,i}^s * u^{s-1}) \right) - \lambda^s A^T \psi(A\, u^{s-1} - m)$$

Gradient Steps



$$u^s_c = u^{s-1}_c - \sum_{i=1}^{N} \bar{k}^s_{c,i} * \phi^s_i \left(\sum_{c \in \{r,g,b\}} (k^s_{c,i} * u^{s-1}) \right) - \lambda^s A^T \psi(A\,u^{s-1} - m)$$

Image Restoration Step by Step - SEM+D8 Model

u^1

Image Restoration Step by Step - SEM+D8 Model

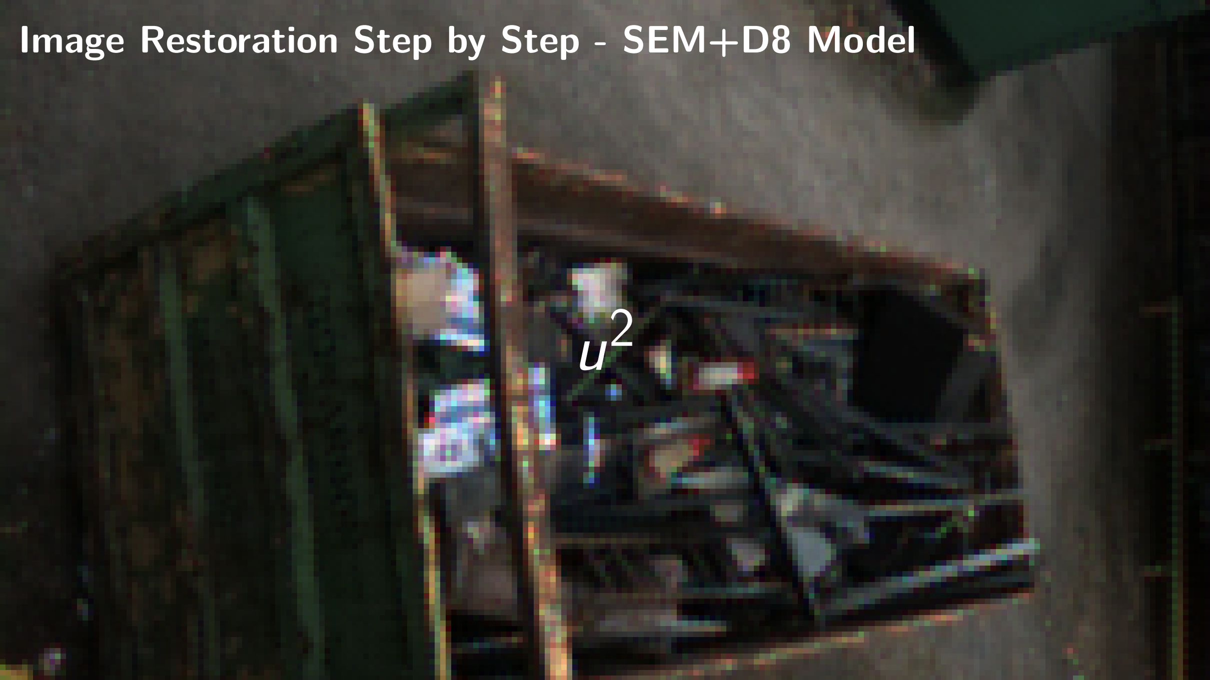


Image Restoration Step by Step - SEM+D8 Model

Image Restoration Step by Step - SEM+D8 Model

Image Restoration Step by Step - SEM+D8 Model

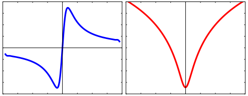
Image Restoration Step by Step - SEM+D8 Model

Image Restoration Step by Step - SEM+D8 Model


Image Restoration Step by Step - SEM+D8 Model
u^8
Reconstruction time: 11ms

Setup for Training

- Data set: 200 training, 200 test, 100 validation images
- Training loss: MSE on linear space images
- Evaluation criteria: PSNR and SSIM on linear space and sRGB images
- Optimization algorithm: LBFGS-B [Byrd et al., 1995]
- Filter sizes
  - Bayer CFA: $5 \times 5 \times 3$ (74 filters)
  - Fujifilm Xtrans CFA: $7 \times 7 \times 3$ (50 filters)
- Filter initialization: Orthonormal DCT basis
- Activation functions: initialized with Student-t function

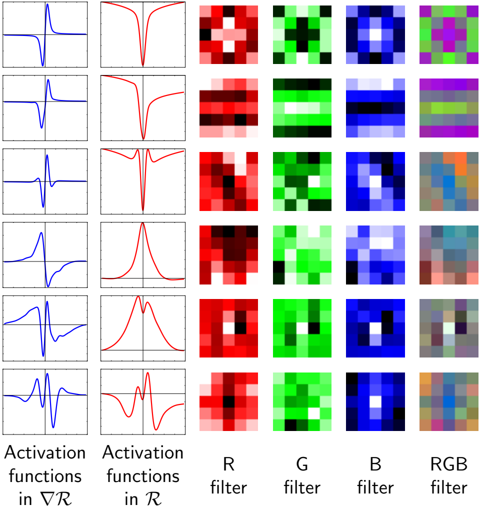


- Initial image: Bilinear interpolation (Bayer), Linear interpolation (Fujifilm Xtrans)
- Intel Core i7 CPU with Nvidia GeForce GTX 980 Ti graphics card

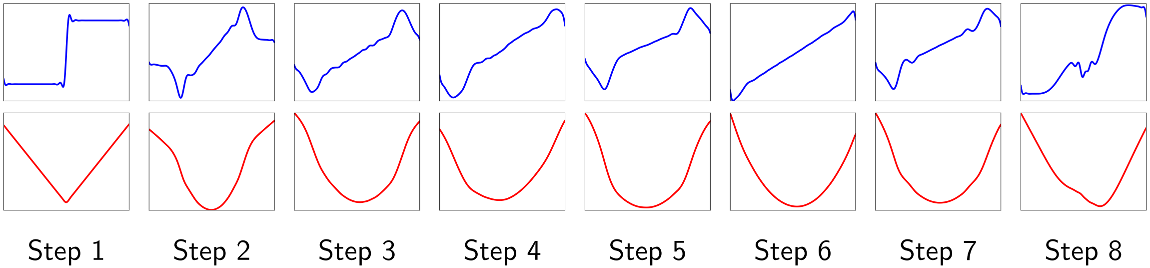
Inspecting the Model

- Some representative examples for activation functions and corresponding kernels
- Radial basis functions

$$\phi_i^s(z, w) = \sum_{j=1}^{M} w_{ij}^s \exp\left(-\frac{(z-\mu_j)^2}{2\sigma^2}\right)$$



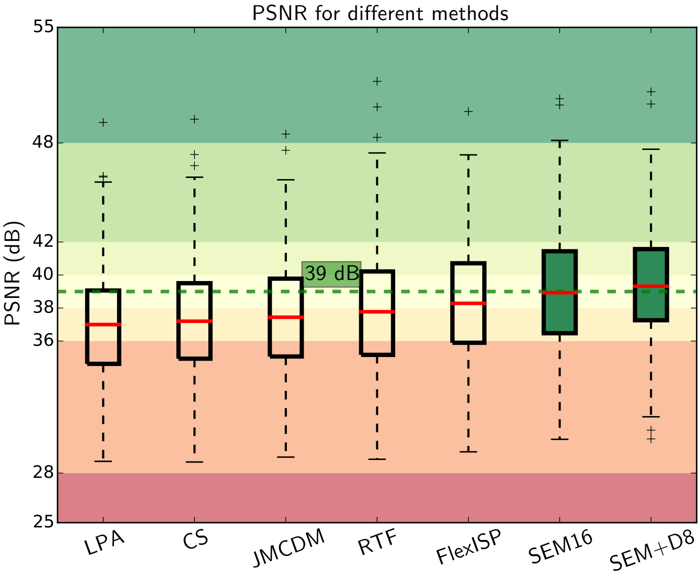
Inspecting the Model - Learned Data Term



Results Bayer CFA

Method	PSNR (linRGB)	PSNR (sRGB)
LPA	37.00	30.86
CS	37.20	31.41
JMCDM	37.44	31.35
RTF	37.77	31.77
FlexISP	38.28	31.76
SEM16	**38.93**	**32.93**
SEM+D8	**39.32**	**33.02**

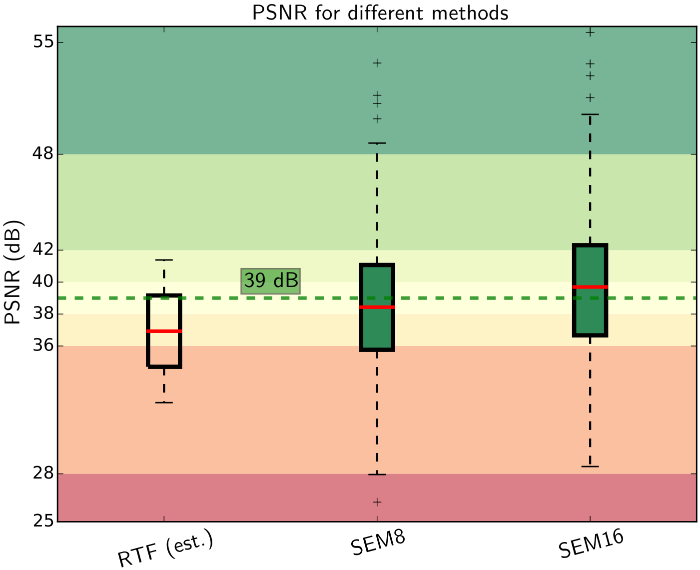
Mean PSNR (test set)



Results
Fujifilm Xtrans CFA

Method	PSNR (linRGB)	PSNR (sRGB)
RTF	36.94	30.56
SEM8	**38.45**	**31.96**
SEM16	**39.60**	**33.09**

Mean PSNR (test set)



Visual Comparisons with Competing Methods

CS

Groundtruth

Visual Comparisons with Competing Methods

JMCDM

Groundtruth

Visual Comparisons with Competing Methods

RTF

Groundtruth

Visual Comparisons with Competing Methods

FlexISP

Groundtruth

Visual Comparisons with Competing Methods

SEM16

Groundtruth

Visual Comparisons SEM+D8 vs. SEM16

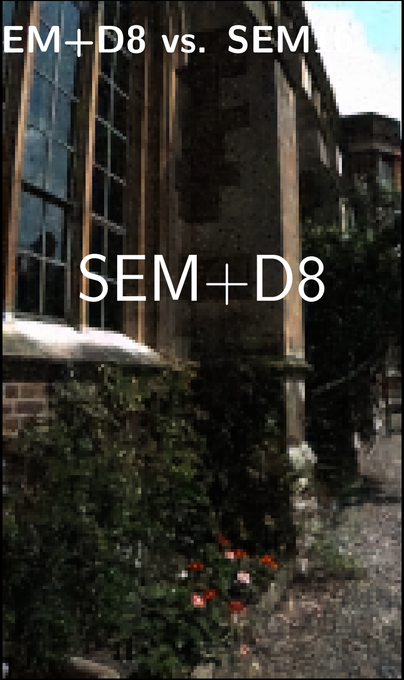
Visual Comparisons SEM+D8 vs. SEM16

Visual Comparisons SEM+D8 vs. SEM16

Visual Comparisons SEM+D8 vs. SEM16

Visual Comparisons SEM+D8 vs. SEM16

Visual Comparisons SEM+D8 vs. SEM



Visual Comparisons SEM+D8 vs. SEM16

Visual Comparisons SEM+D8 vs. SEM16

SEM+D8

Test Image Taken with Sony A6000 (6MP)


Test Image Taken with Sony A6000 (6MP)
SEM16

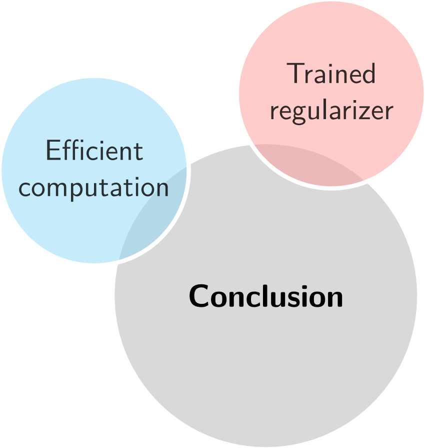
Test Image Taken with Sony A6000 (6MP)

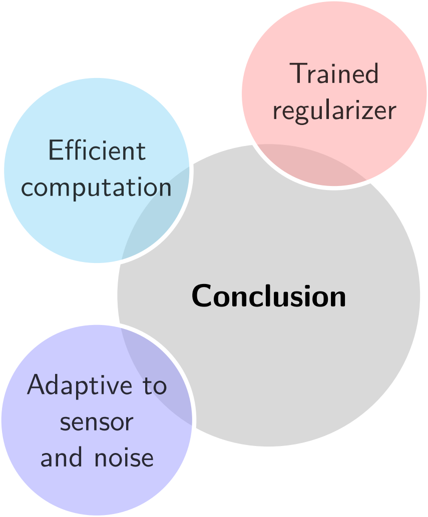
Test Image Taken with Sony A6000 (6MP)

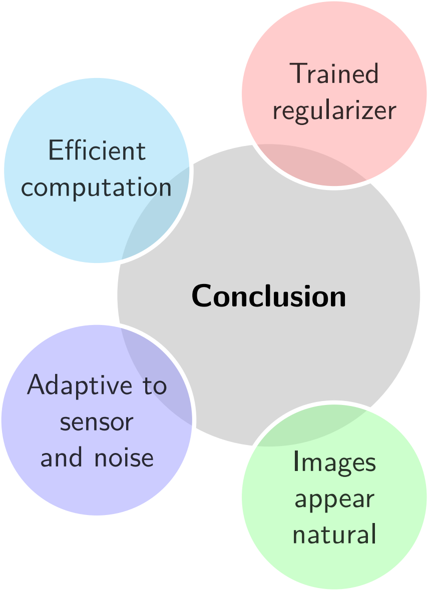
Conclusion

Trained regularizer

Conclusion


Trained regularizer
Efficient computation
Conclusion


Trained regularizer
Efficient computation
Conclusion
Adaptive to sensor and noise


Trained regularizer
Efficient computation
Conclusion
Adaptive to sensor and noise
Images appear natural

Conclusion

- Trained regularizer
- State-of-the-art results
- Images appear natural
- Adaptive to sensor and noise
- Efficient computation

Acknowledgements

- Thanks to my colleagues Kerstin Hammernik, Patrick Knöbelreiter, Erich Kobler

- Thanks to my supervisor Prof. Thomas Pock

- Thanks to the Austrian Science Fund (FWF) for support under the START project BIVISION, No. Y729.
- And finally ...

Thank you!