Joint Content-Adaptive Dictionary Learning and Sparse Selective Extrapolation for Cross-Spectral Image Reconstruction

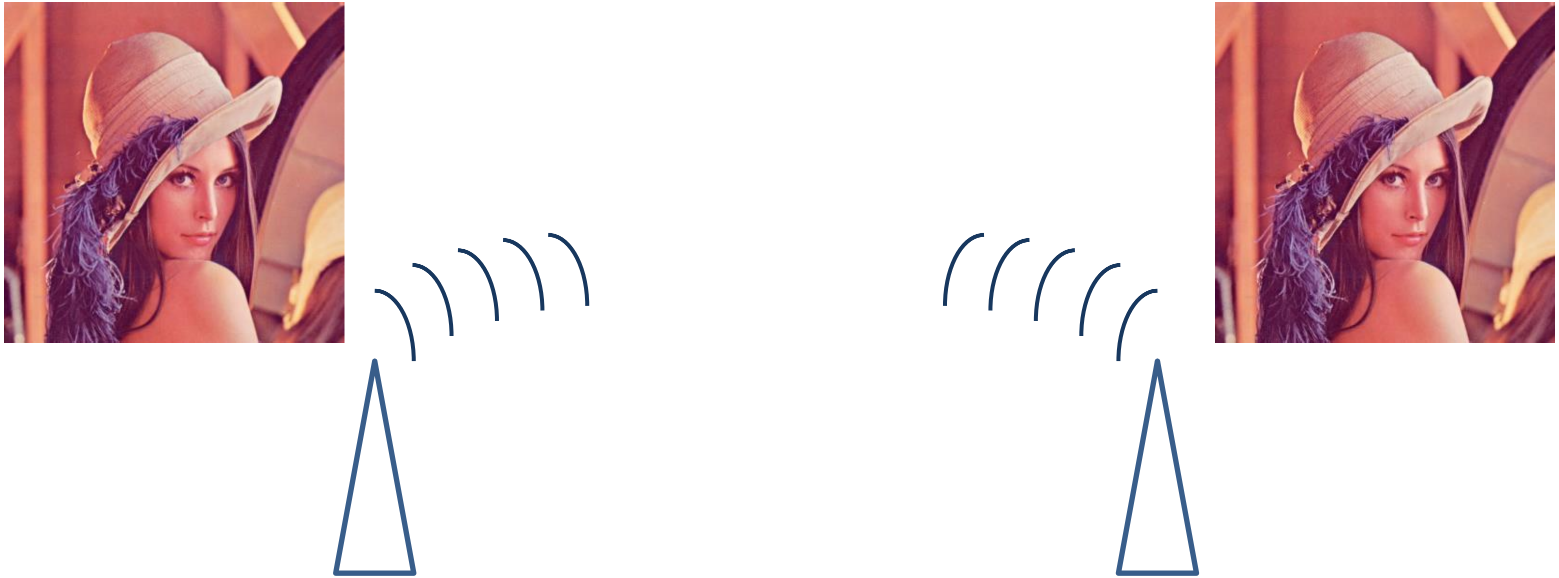
Nils Genser, Jürgen Seiler, and André Kaup
nils.genser@fau.de
Chair of Multimedia Communications and Signal Processing

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

(TEC-09 -- Restoration and Enhancement II)

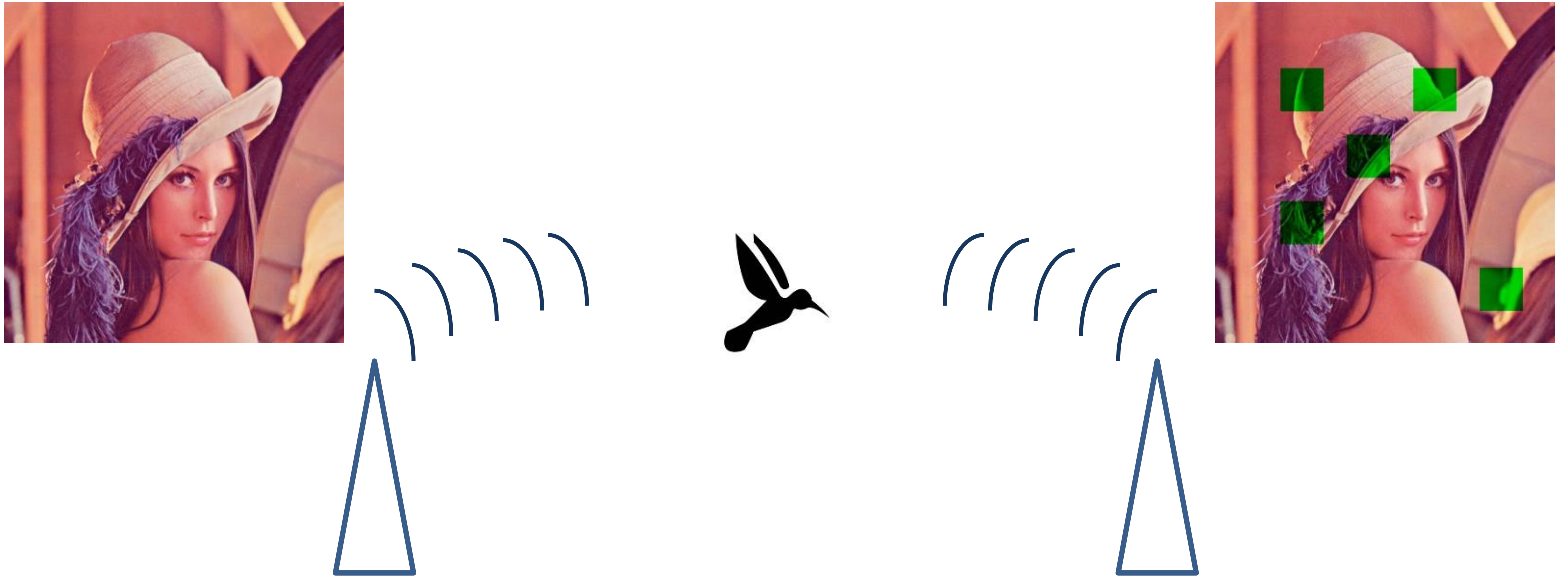
Use Case: Error Concealment

Wireless Video Communication

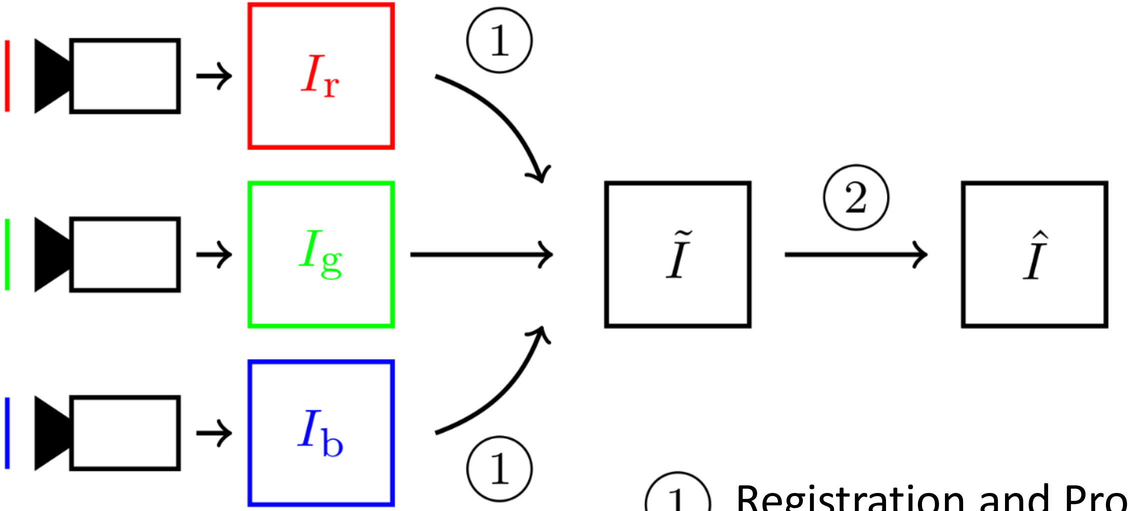


Use Case: Error Concealment

Wireless Video Communication → Transmission errors



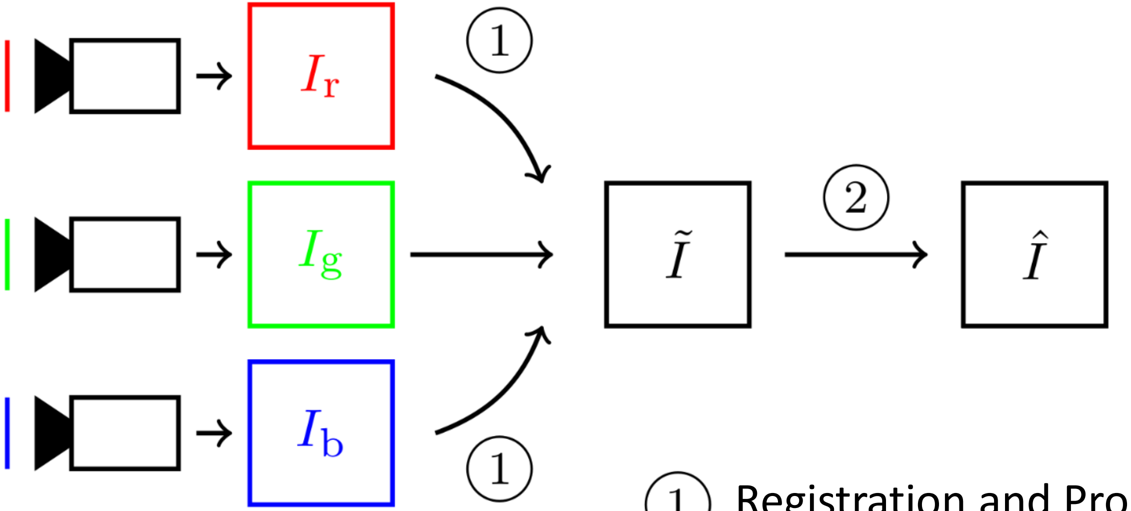
Use Case: Multi-Color Multi-Camera Setups



1. Registration and Projection
2. Color Channel Reconstruction

[1] Kiesel et al., "Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression," in Proc. Electronic Imaging - Image Sensors and Imaging Systems Conference, Jan. 2018

Use Case: Multi-Color Multi-Camera Setups



1. Registration and Projection
2. Color Channel Reconstruction

[1] Kiesel et al., "Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression," in Proc. Electronic Imaging - Image Sensors and Imaging Systems Conference, Jan. 2018

Multi-Color Multi-Camera Setups

Registered and projected image

?

Reconstructed image

Outline

- Motivation
- State-of-the-Art Techniques
- Novel Joint Content-Adaptive Dictionary Learning and Sparse Selective Extrapolation
- Evaluation
- Conclusion and Reference Implementation

State-of-the-Art Reconstruction I

Exploit spatial similarity:

- Very well investigated problem in inpainting and error concealment
- Popular approach: Total Variation (TV) **[2]**
- Better (in terms of quality): Frequency Selective Extrapolation (FSE) **[3]**

Sparse model: $\boldsymbol{g} = \sum_{k \in \mathcal{K}} \hat{c}_k \cdot \boldsymbol{\varphi}_k$

with estimated coefficients $\hat{c}_k$

and Fourier basis functions $\boldsymbol{\varphi}_k$ in set $\mathcal{K}$

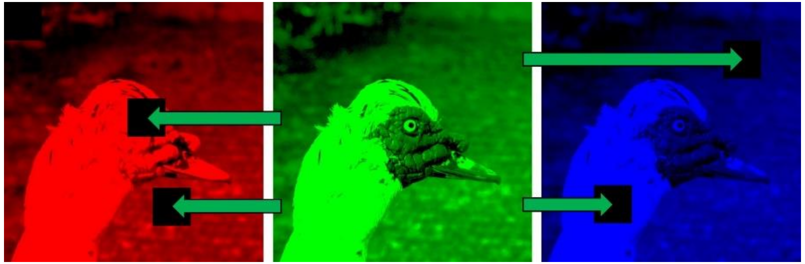
[2] Dahl et al., "Algorithms and software for total variation image reconstruction via first-order methods," in Numerical Algorithms, Jul. 2009

[3] Genser et al., "Spectrally constrained frequency selective extrapolation for rapid image error concealment," in Proc. International Conference on Systems, Signals and Image Processing, Jun. 2018

State-of-the-Art Reconstruction II

Exploit spectral similarity:

- Fundamental idea: Color components highly correlated
- Model for predicting missing information (using existing components)
- Color Similarity-Based Local Median Filtering (CSLMF) **[1]**
  - Pixel-wise processing of missing samples
  - Reconstruction → median-filtered local neighborhood
  - Reference channel supports candidates



[1] Kiesel et al., “Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression,” in Proc. Electronic Imaging - Image Sensors and Imaging Systems Conference, Jan. 2018

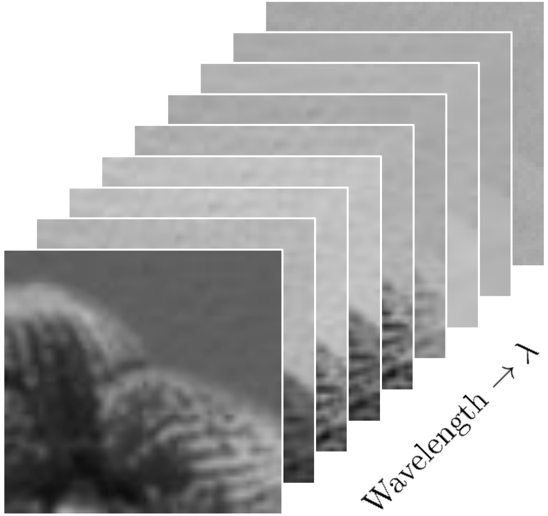
Joint Content-Adaptive Dictionary Learning and Sparse Selective Extrapolation (CASE)

THE NOVEL APPROACH

Motivation

Idea:

- Image components typically highly correlated
- ➢ Undistorted channels as reference
- Block-based processing of distorted image

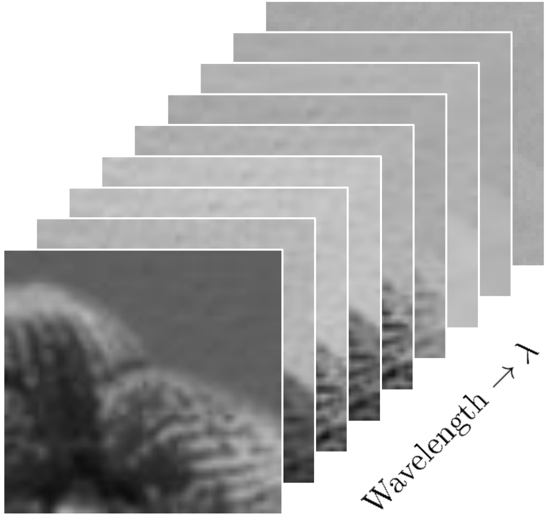


[4] Seiler et al., “A fast algorithm for selective signal extrapolation with arbitrary basis functions,” in EURASIP Journal on Advances in Signal Processing, vol. 2011, pp. 1–10, Jan. 2011.

Motivation

Idea:

- Image components typically highly correlated
- ➢ Undistorted channels as reference
- Block-based processing of distorted image



Goal: Derive dictionary $\mathcal{D}$ of basis functions φ_d

Model: $\boldsymbol{g} = \sum_{d \in \mathcal{D}} \hat{c}_d \cdot \varphi_d$

Tasks:

- Create content-adaptive dictionary from undistorted channels
- Estimate coefficients $\hat{c}_d$ using **[4]**

[4] Seiler et al., "A fast algorithm for selective signal extrapolation with arbitrary basis functions," in EURASIP Journal on Advances in Signal Processing, vol. 2011, pp. 1–10, Jan. 2011.

Best Reference Selection

References $\mathcal{R}$ → **Best reference selection** → r_{best}, $\mathcal{R}$ → Redundancy reduction → $\mathcal{R}_\Delta$ → Singular value decomposition → Denoising → Singular value composition → Dictionary $\mathcal{D}$

Preparation: Create set $\mathcal{R}$ of references $\boldsymbol{r}_k$

$$\mathcal{R} = \bigcup_{k \in \mathcal{K}} \boldsymbol{r}_k$$

Correlation analysis:

$$p_k = \frac{||(\tilde{\boldsymbol{d}} - \overline{d}) \cdot (\tilde{\boldsymbol{r}}_k - \overline{r}_k)||_1}{||\tilde{\boldsymbol{d}} - \overline{d}||_2 \cdot ||\tilde{\boldsymbol{r}}_k - \overline{r}_k||_2} \in [0, 1]$$

Best reference (highest p_k): $\tilde{\boldsymbol{r}}_{\text{best}}$

$\boldsymbol{d}$: Distorted signal as vector

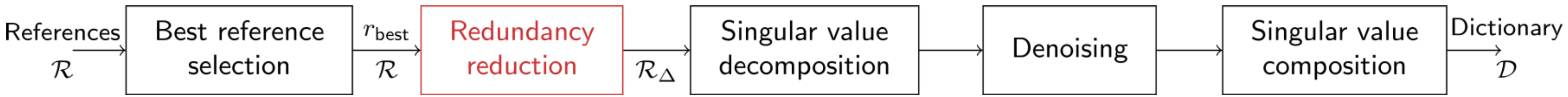
$\tilde{\boldsymbol{d}}$: Valid samples of $\boldsymbol{d}$

$\boldsymbol{r}_k$: k-th reference

$\tilde{\boldsymbol{r}}_k$: Reference according to $\tilde{\boldsymbol{d}}$

k : Fully conserved references with $k \in \mathcal{K}$

Redundancy Reduction



Delta reference list:

$$\mathcal{R}_\Delta = \bigcup_{k \in \mathcal{K}} \boldsymbol{r}_{\Delta\, k}$$

with zero-mean difference references

$$\boldsymbol{r}_{\Delta\, k} = \begin{cases} \boldsymbol{r}_k - \overline{r}_k, & k = k_{\text{best}} \\ \boldsymbol{r}_k - \overline{r}_k - \boldsymbol{r}_{\text{best}} + \overline{r}_{\text{best}}, & \text{else} \end{cases}$$

Singular Value Decomposition

References $\mathcal{R}$ → Best reference selection → r_{best}, $\mathcal{R}$ → Redundancy reduction → $\mathcal{R}_\Delta$ → Singular value decomposition → Denoising → Singular value composition → Dictionary $\mathcal{D}$

Delta reference list:

$$\mathcal{R}_\Delta = \bigcup_{k \in \mathcal{K}} \boldsymbol{r}_{\Delta\, k}$$

with zero-mean difference references

$$\boldsymbol{r}_{\Delta\, k} = \begin{cases} \boldsymbol{r}_k - \overline{r}_k, & k = k_{\text{best}} \\ \boldsymbol{r}_k - \overline{r}_k - \boldsymbol{r}_{\text{best}} + \overline{r}_{\text{best}}, & \text{else} \end{cases}$$

Singular value decomposition:

$$\boldsymbol{r}_{\Delta\, k} = \boldsymbol{U}_k \boldsymbol{\Sigma}_k \boldsymbol{V}_k^*$$

New singular value matrices:

$$\boldsymbol{\Sigma}_{k\, b} = \begin{bmatrix} s_b & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

with singular values s_b

and number of diagonal entries b

Denoising

References $\mathcal{R}$ → Best reference selection → r_{best} $\mathcal{R}$ → Redundancy reduction → $\mathcal{R}_\Delta$ → Singular value decomposition → **Denoising** → Singular value composition → Dictionary $\mathcal{D}$

Delta reference list:

$$\mathcal{R}_\Delta = \bigcup_{k \in \mathcal{K}} \boldsymbol{r}_{\Delta\, k}$$

with zero-mean difference references

$$\boldsymbol{r}_{\Delta\, k} = \begin{cases} \boldsymbol{r}_k - \overline{r}_k, & k = k_{\text{best}} \\ \boldsymbol{r}_k - \overline{r}_k - \boldsymbol{r}_{\text{best}} + \overline{r}_{\text{best}}, & \text{else} \end{cases}$$

Omit new singular value matrices, with $s_b < 0.01$

Singular value decomposition:

$$\boldsymbol{r}_{\Delta\, k} = \boldsymbol{U}_k \boldsymbol{\Sigma}_k \boldsymbol{V}_k^*$$

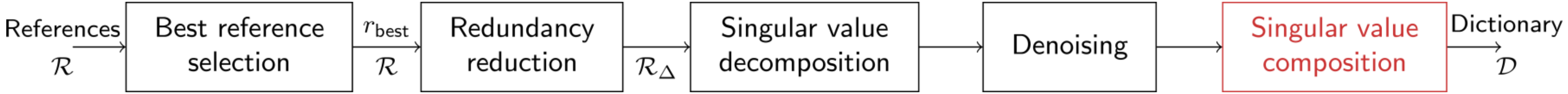
New singular value matrices:

$$\boldsymbol{\Sigma}_{k\, b} = \begin{bmatrix} s_b & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

with singular values s_b

and number of diagonal entries b

Singular Value Composition



Dictionary:

$$\mathcal{D} = \varphi_{\text{DC}} \bigcup_{k \in \mathcal{K}} \bigcup_{b \in \mathcal{B}_k} \varphi_{k\,b}$$

with DC basis function: $\varphi_{\text{DC}} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{M \times M}$

and reference based basis functions: $\varphi_{k\,b} = \boldsymbol{U}_k \boldsymbol{\Sigma}_{k\,b} \boldsymbol{V}_k^*$

Model Generation

Model: $\boldsymbol{g} = \sum_{d \in \mathcal{D}} \hat{c}_d \cdot \boldsymbol{\varphi}_d$

Estimate coefficients using **[4]**

Estimated prediction:

$$\boldsymbol{p}_i = \begin{cases} \boldsymbol{d}_i\,, & i \in \mathcal{V} \\ \boldsymbol{g}_i\,, & \text{else} \end{cases}\,, \quad \forall i \in \mathcal{A}$$

i : Pixel position

$\mathcal{V}$: Set of valid indices

$\mathcal{A}$: Set of all indices

Model Generation

Model: $\boldsymbol{g} = \sum_{d \in \mathcal{D}} \hat{c}_d \cdot \boldsymbol{\varphi}_d$

Estimate coefficients using **[4]**

Estimated prediction:

$$\boldsymbol{p}_i = \begin{cases} \boldsymbol{d}_i \,, & i \in \mathcal{V} \\ \boldsymbol{g}_i \,, & \text{else} \end{cases} \,, \quad \forall i \in \mathcal{A}$$

i : Pixel position

$\mathcal{V}$: Set of valid indices

$\mathcal{A}$: Set of all indices

Process all distorted blocks
➢ Reconstructed image

Evaluation

PERFORMANCE ANALYSIS

Objective Evaluation I

Novel CASE:

- Training: Kodak data set, block losses
- Evaluation: TECNICK/Middlebury data sets

	Spatial		Cross-spectral		
	TV [2]	FSE [3]	CSLMF [1]	JRR [5]	Novel CASE
TECNICK - Pattern block					
PSNR	28.3 dB	30.2 dB	33.2 dB	38.6 dB	**39.1 dB**
SSIM	0.951	0.960	0.976	0.992	**0.993**
TECNICK - Pattern line					
PSNR	23.2 dB	24.5 dB	30.3 dB	31.6 dB	**32.7 dB**
SSIM	0.871	0.886	0.956	0.975	**0.977**
TECNICK - Pattern block-line					
PSNR	26.6 dB	28.6 dB	32.9 dB	36.5 dB	**36.9 dB**
SSIM	0.941	0.954	0.976	0.990	**0.991**
TECNICK - Pattern random					
PSNR	31.8 dB	33.7 dB	34.4 dB	40.8 dB	**41.9 dB**
SSIM	0.974	0.979	0.980	0.995	**0.996**
Middlebury - Disparity maps					
PSNR	28.6 dB	30.2 dB	33.4 dB	33.7 dB	**35.2 dB**
SSIM	0.946	0.967	0.981	0.985	**0.988**

[1] Kiesel et al., "Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression," in Proc. Electronic Imaging - Image Sensors and Imaging Systems Conference, Jan. 2018
[2] Dahl et al., "Algorithms and software for total variation image reconstruction via first-order methods," in Numerical Algorithms, Jul. 2009
[3] Genser et al., "Spectrally constrained frequency selective extrapolation for rapid image error concealment," in Proc. International Conference on Systems, Signals and Image Processing, Jun. 2018
[5] Genser et al., "Joint Regression Modeling and Sparse Spatial Refinement for High-Quality Reconstruction of Distorted Color Images," in Proc. IEEE International Conference on Image Processing (ICIP), Sep. 2019

Objective Evaluation I

Novel CASE:

- Training: Kodak data set, block losses
- Evaluation: TECNICK/Middlebury data sets
- Outperforms state-of-the-art :
  - error concealment algorithms
  - color reconstruction approaches
- ➢ Suited for all kind of error patterns

	Spatial		Cross-spectral		
	TV [2]	FSE [3]	CSLMF [1]	JRR [5]	Novel CASE
TECNICK - Pattern block					
PSNR	28.3 dB	30.2 dB	33.2 dB	38.6 dB	**39.1 dB**
SSIM	0.951	0.960	0.976	0.992	**0.993**
TECNICK - Pattern line					
PSNR	23.2 dB	24.5 dB	30.3 dB	31.6 dB	**32.7 dB**
SSIM	0.871	0.886	0.956	0.975	**0.977**
TECNICK - Pattern block-line					
PSNR	26.6 dB	28.6 dB	32.9 dB	36.5 dB	**36.9 dB**
SSIM	0.941	0.954	0.976	0.990	**0.991**
TECNICK - Pattern random					
PSNR	31.8 dB	33.7 dB	34.4 dB	40.8 dB	**41.9 dB**
SSIM	0.974	0.979	0.980	0.995	**0.996**
Middlebury - Disparity maps					
PSNR	28.6 dB	30.2 dB	33.4 dB	33.7 dB	**35.2 dB**
SSIM	0.946	0.967	0.981	0.985	**0.988**

[1] Kiesel et al., “Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression,” in Proc. Electronic Imaging - Image Sensors and Imaging Systems Conference, Jan. 2018

[2] Dahl et al., “Algorithms and software for total variation image reconstruction via first-order methods,” in Numerical Algorithms, Jul. 2009

[3] Genser et al., “Spectrally constrained frequency selective extrapolation for rapid image error concealment,” in Proc. International Conference on Systems, Signals and Image Processing, Jun. 2018

[5] Genser et al., “Joint Regression Modeling and Sparse Spatial Refinement for High-Quality Reconstruction of Distorted Color Images,” in Proc. IEEE International Conference on Image Processing (ICIP), Sep. 2019

Objective Evaluation II

Computation time:

- Implementation: MATLAB
- Averaged over 100 executions
- ➢ Less complex than state of the art

	TV [2]	FSE [3]	CSLMF [1]	JRR [5]	Novel CASE
TECNICK - Pattern block					
Time	19 s	103 s	257 s	104 s	**13 s**
TECNICK - Pattern line					
Time	35 s	173 s	449 s	188 s	**33 s**
TECNICK - Pattern block-line					
Time	29 s	89 s	231 s	93 s	**18 s**
TECNICK - Pattern random					
Time	31 s	92 s	212 s	94 s	**20 s**
Middlebury - Disparity maps					
Time	51 s	221 s	1487 s	225 s	**49 s**

[1] Kiesel et al., "Color Channel Reconstruction for Multi-Color Multi-View Images Using Disparity and Color Similarity-based Local Linear Regression," in Proc. Electronic Imaging - Image Sensors and Imaging Systems Conference, Jan. 2018
[2] Dahl et al., "Algorithms and software for total variation image reconstruction via first-order methods," in Numerical Algorithms, Jul. 2009
[3] Genser et al., "Spectrally constrained frequency selective extrapolation for rapid image error concealment," in Proc. International Conference on Systems, Signals and Image Processing, Jun. 2018
[5] Genser et al., "Joint Regression Modeling and Sparse Spatial Refinement for High-Quality Reconstruction of Distorted Color Images," in Proc. IEEE International Conference on Image Processing (ICIP), Sep. 2019

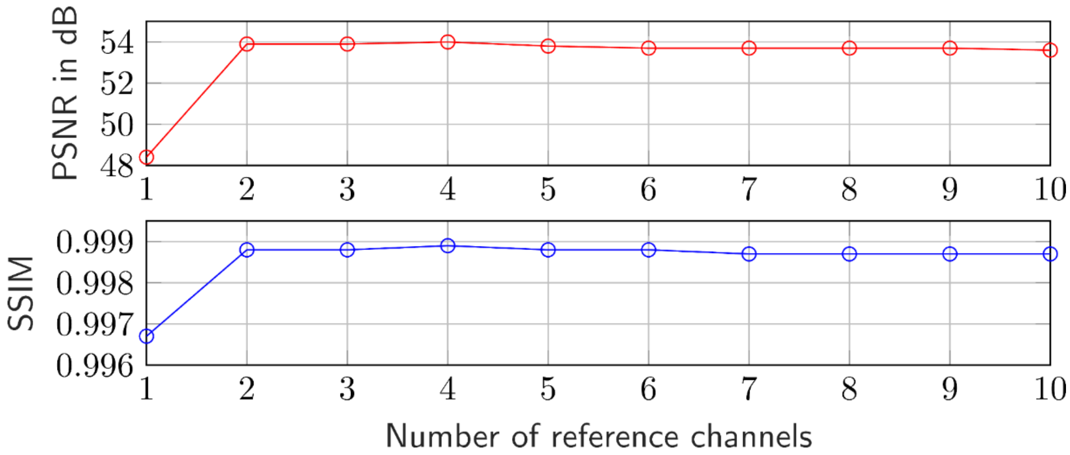
Objective Evaluation III

Robustness:

- Evaluation: Multi-spectral CAVE data set (400 to 700 nm)
- Image at 550 nm distorted
- Combination of block, random, and line losses
- References: Spectrally varying, adjacent first

➢ Multiple references beneficial

➢ Little influence for bad references



Visual Evaluation I

Ground truth

Distorted

Visual Evaluation II

Ground truth

State-of-the-art TV **[2]**

Visual Evaluation III

Ground truth

State-of-the-art CSLMF **[1]**

Visual Evaluation IV

Ground truth

Proposed CASE

Conclusion

AND REFERENCE IMPLEMENTATION

Conclusion

- Distorted color/multispectral images:
  E.g., in video coding or multi-camera setups

Conclusion

- Distorted color/multispectral images:
  E.g., in video coding or multi-camera setups
- State-of-the-art techniques
  - Either cross-component or spatial reconstruction
  - Hybrid models for only one reference component
  - Insufficient quality

Conclusion

- Distorted color/multispectral images:
  E.g., in video coding or multi-camera setups

- State-of-the-art techniques
  - Either cross-component or spatial reconstruction
  - Hybrid models for only one reference component
  - Insufficient quality

- Novel CASE algorithm
  - Spatial and cross-color reconstruction
  - Arbitrary number of reference channels possible
  - Outstanding visual and objective quality

Reference Implementation

CASE framework :

- Color component reconstruction using multiple reference channels
- *https://gitlab.lms.tf.fau.de/lms/case*