Envision Future COMPUTING
Computing for the FUTURE

Analysis and Synthesis Sparse Representation Models for Image Modeling

Shuhang GU

Dept. of Computing

The Hong Kong Polytechnic University

Outline

- Introduction
  - Image Restoration and Enhancement
  - Synthesis & Analysis Sparsity Models
- Convolutional sparse coding for image super-resolution
  - Convolutional Sparse Coding v.s. Sparse Coding
  - The Proposed Method
- Guided Image Enhancement via Weighted Analysis Sparsity Model
  - Dependency Modeling for Guided Enhancement
  - Learning dynamic guidance for guided depth enhancement
- Ongoing and Future Works
  - Image Separation without Training Data
  - Image Restoration with Deep Denoisers
  - Optimization Inspired Network Structure Design

Introduction

Image restoration and enhancement problems

Image Denoising

$$y = x + n$$

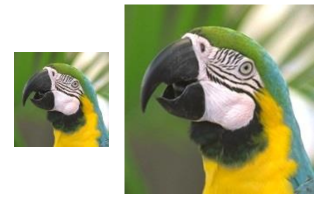
Image Deconvolution

$$y = k \otimes x + n$$

Contrast Enhancement

?

Image Super-resolution



$$y = D(k \otimes x) + n$$

Image Inpainting

$$y = M \odot x + n$$

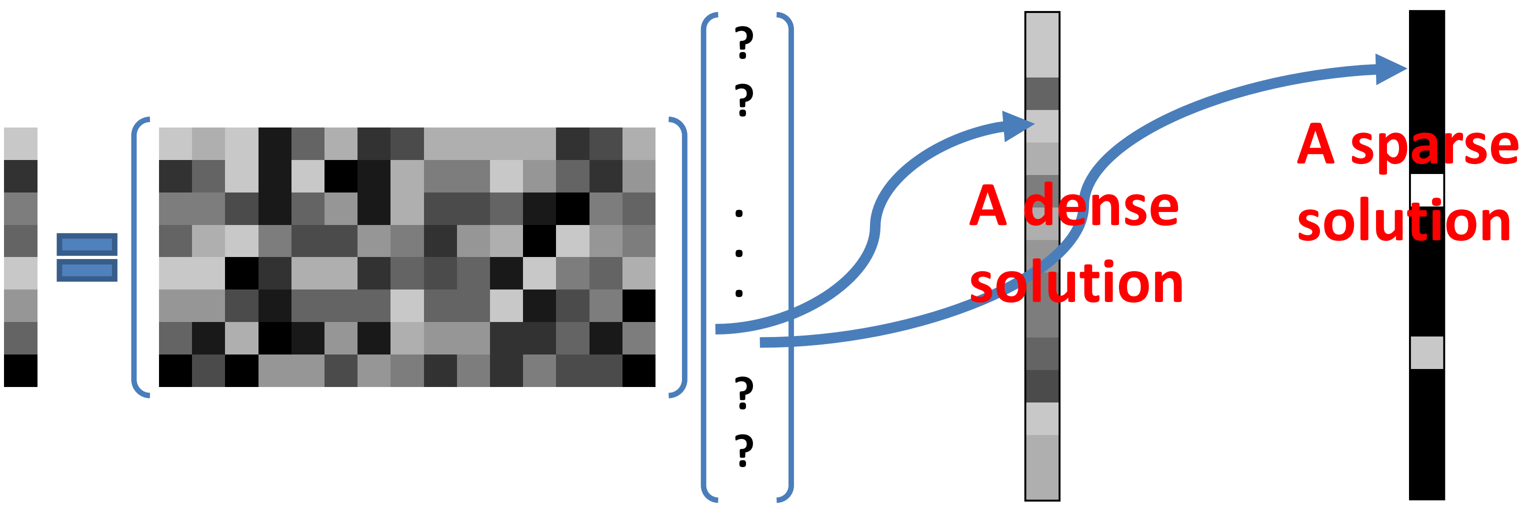
…

Synthesis & Analysis sparsity models

- Synthesis representation models

Synthesis based sparse representation model assumes that a signal x can be represented as a linear combination of a small number of atoms chosen out of a dictionary D:

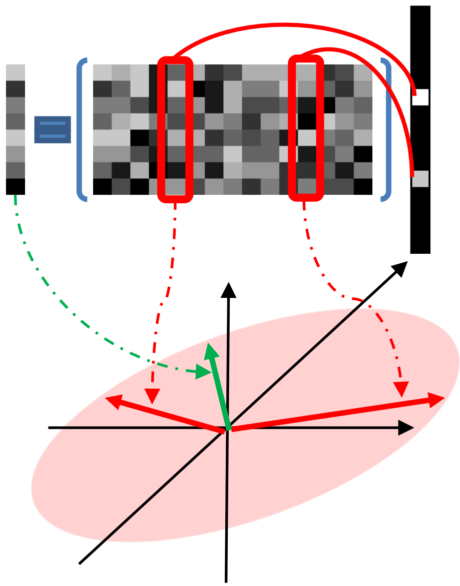
$$x = D\alpha,\ s.t.\ \|\alpha\|_0 < \varepsilon$$



Synthesis & Analysis sparsity models

- Synthesis representation models

Synthesis based sparse representation model assumes that a signal x can be represented as a linear combination of a small number of atoms chosen out of a dictionary D:

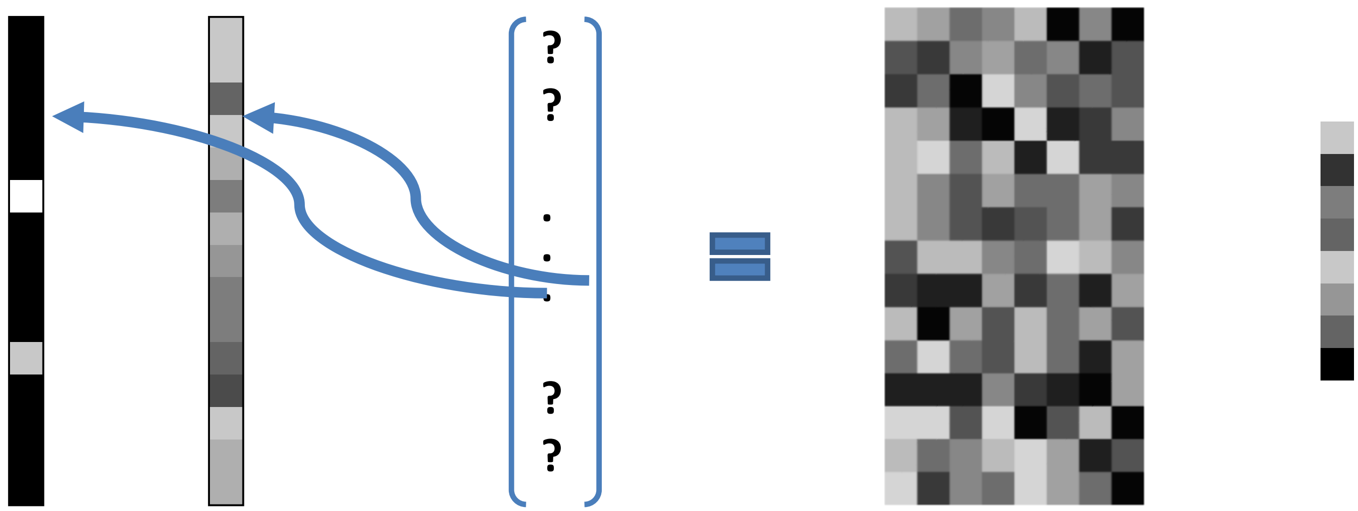


Synthesis & Analysis sparsity models

- Analysis representation models

Analysis model generate representation coefficients by a simple multiplication operation, and assumes the coefficients are sparse:

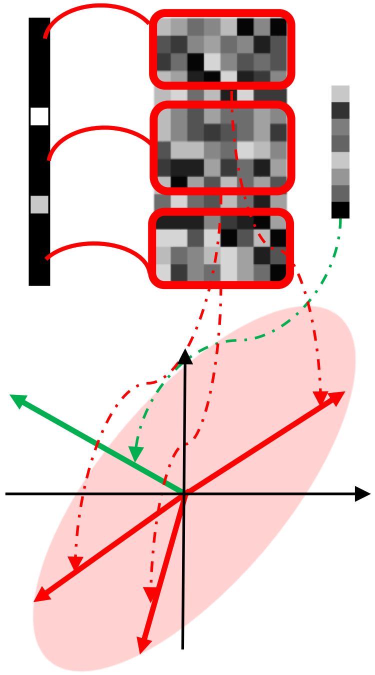
$$||Px||_0 < \varepsilon$$



Synthesis & Analysis sparsity models

- Analysis representation models

Analysis model generate representation coefficients by a simple multiplication operation, and assumes the coefficients are sparse:



Synthesis & Analysis sparsity models

Synthesis model

$$min_{\alpha} \frac{1}{2} \|y - D\alpha\|_F^2 + \psi(\alpha)$$

$$x = D\alpha$$

- Representative methods

KSVD, BM3D, LSSC, NCSR, et. al.

- Pros

- Synthesis model can be more sparse

- Cons

- Patch prior modeling needs aggregation

- Time consuming

Analysis model

$$min_{x} \frac{1}{2} \|y - x\|_F^2 + \phi(Px)$$

- Representative methods

TV, wavelet methods, FRAME, FOE, CSF, TRD et. al.

- Pros

- Patch divide free

- Efficient in the inference phase

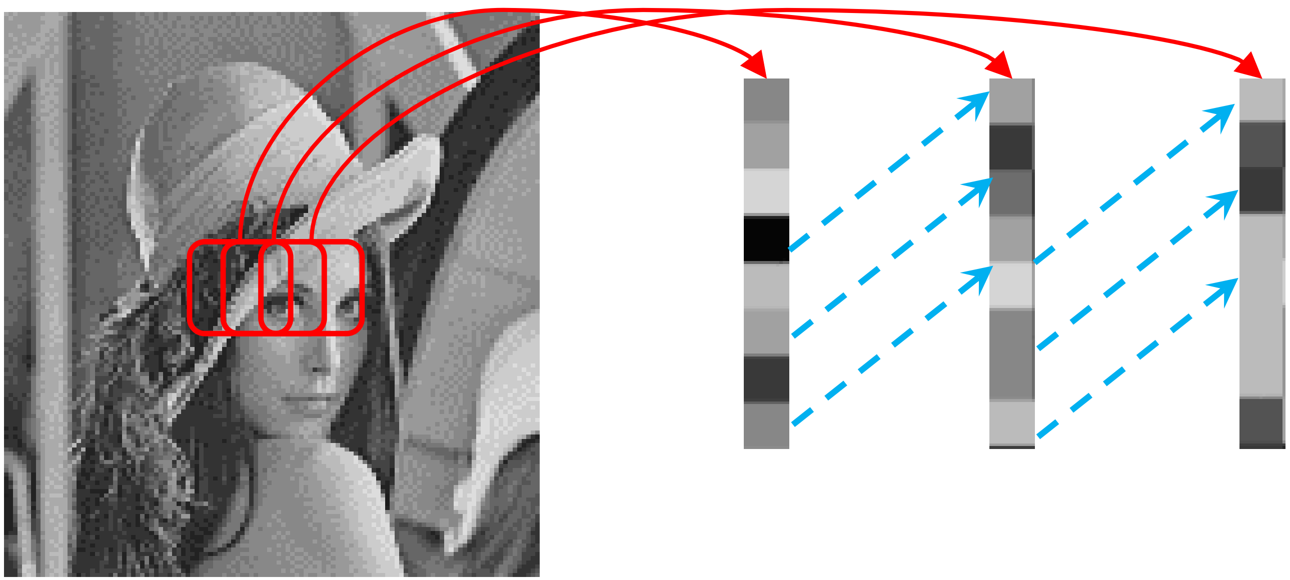
- Cons

- Not as sparse as synthesis model, limited capacity in modeling image prior.

Convolutional sparse coding for image super-resolution

Convolutional Sparse Coding v.s. Sparse Coding

Consistency constraint



Convolutional Sparse Coding v.s. Sparse Coding

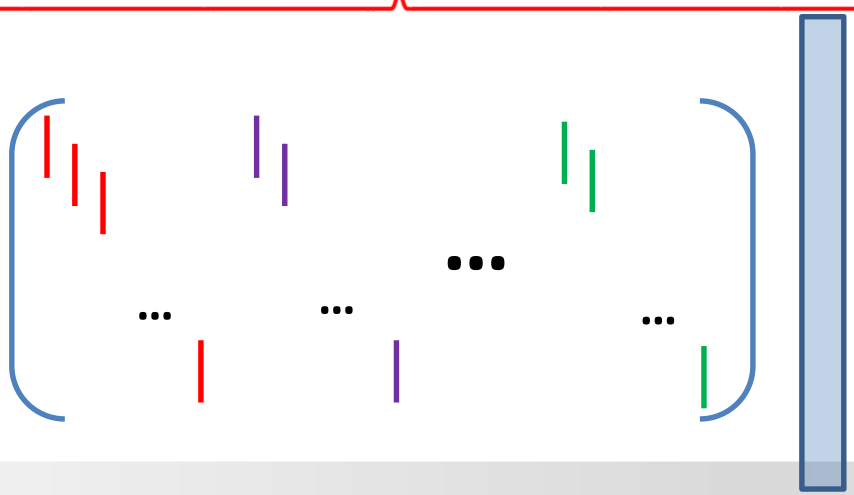
Sparse coding

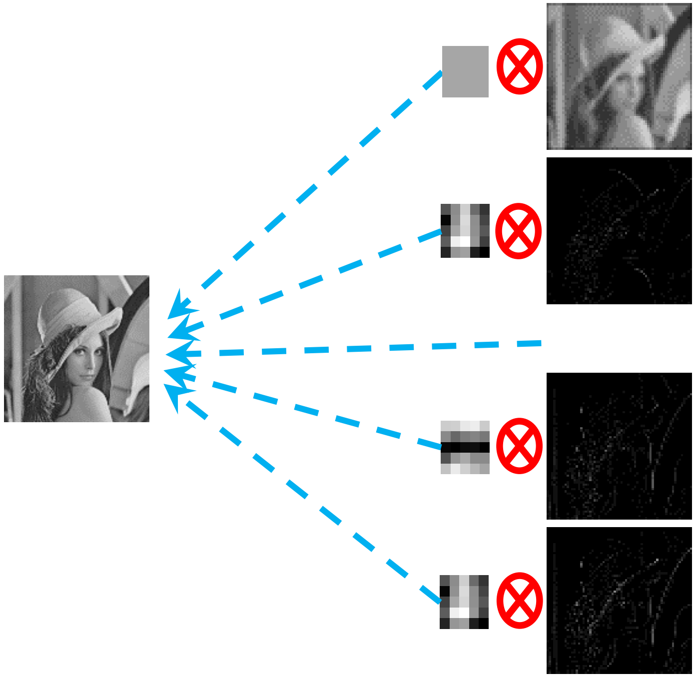
$$min_{\alpha}||y-D\alpha||_F^2+\phi(\alpha)$$

Convolutional sparse coding

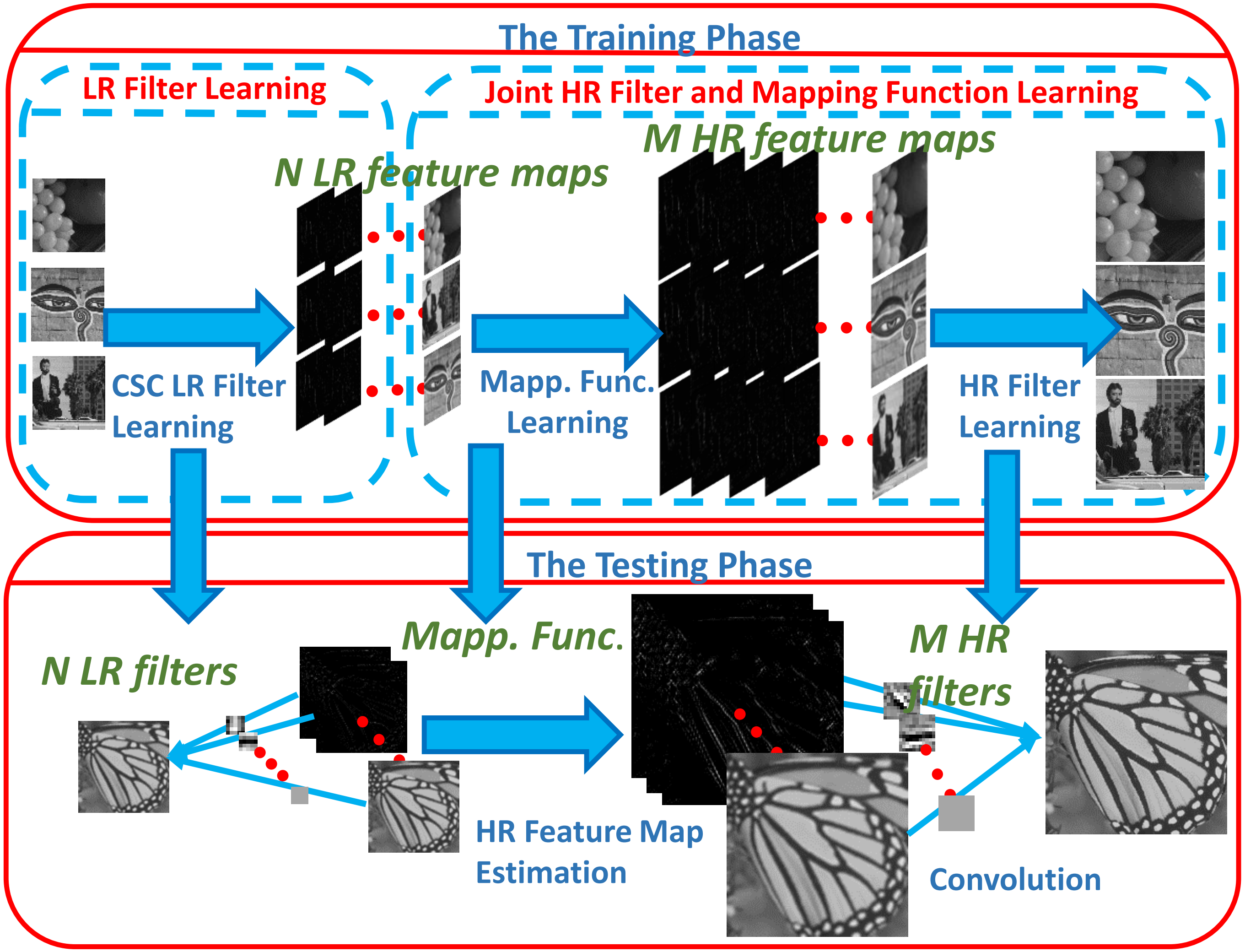
$$min_{\boldsymbol{z}}||\boldsymbol{Y}-\underline{\sum \boldsymbol{f}_i \otimes \boldsymbol{z}_i}||_F^2+\sum \varphi(\boldsymbol{z}_i)$$

Matrix Form

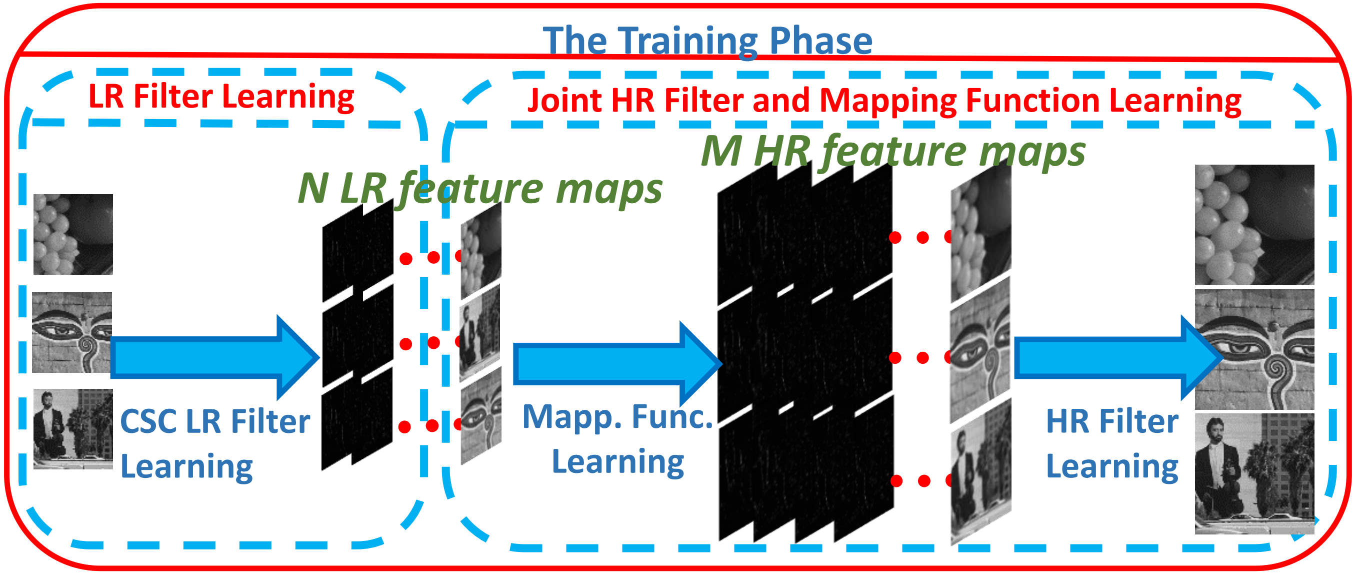




Convolutional sparse coding for image SR



Convolutional sparse coding for image SR

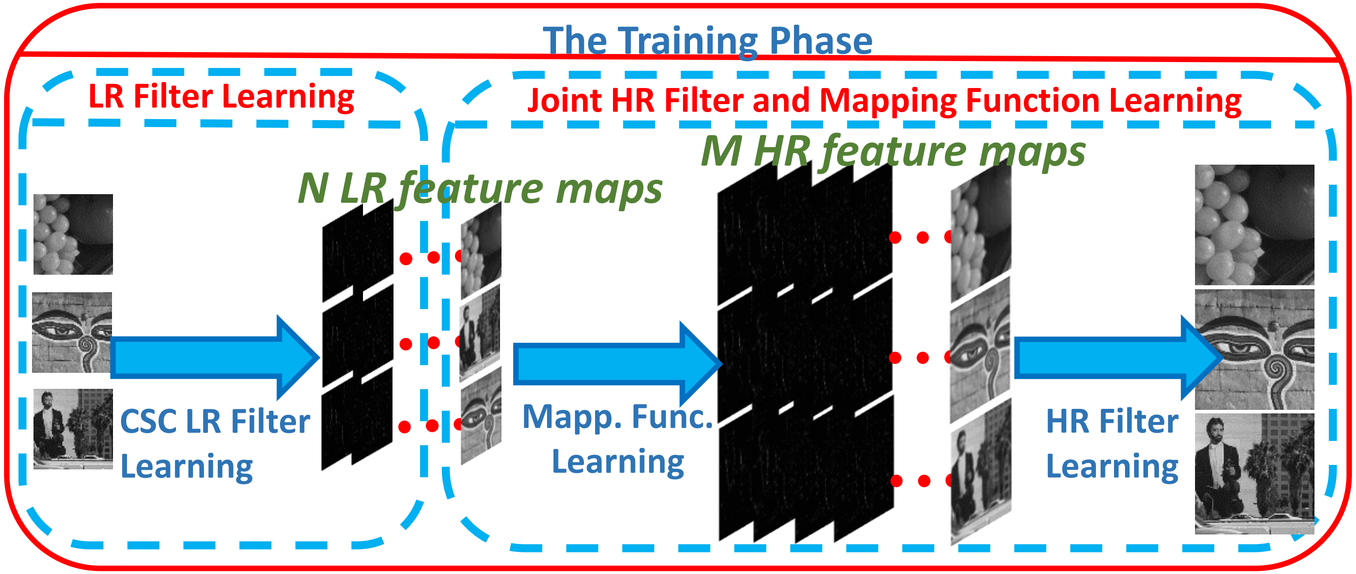


– LR filter training

$$\min_{Z,f} \|Y - \sum_{i=1}^{N} f_i^l \otimes Z_i^l \|_F^2 + \lambda \sum_{i=1}^{N} \| Z_i^l \|_1$$

$$s.t. \ \| f_i^l \|_F^2 \leq 1$$

Convolutional sparse coding for image SR

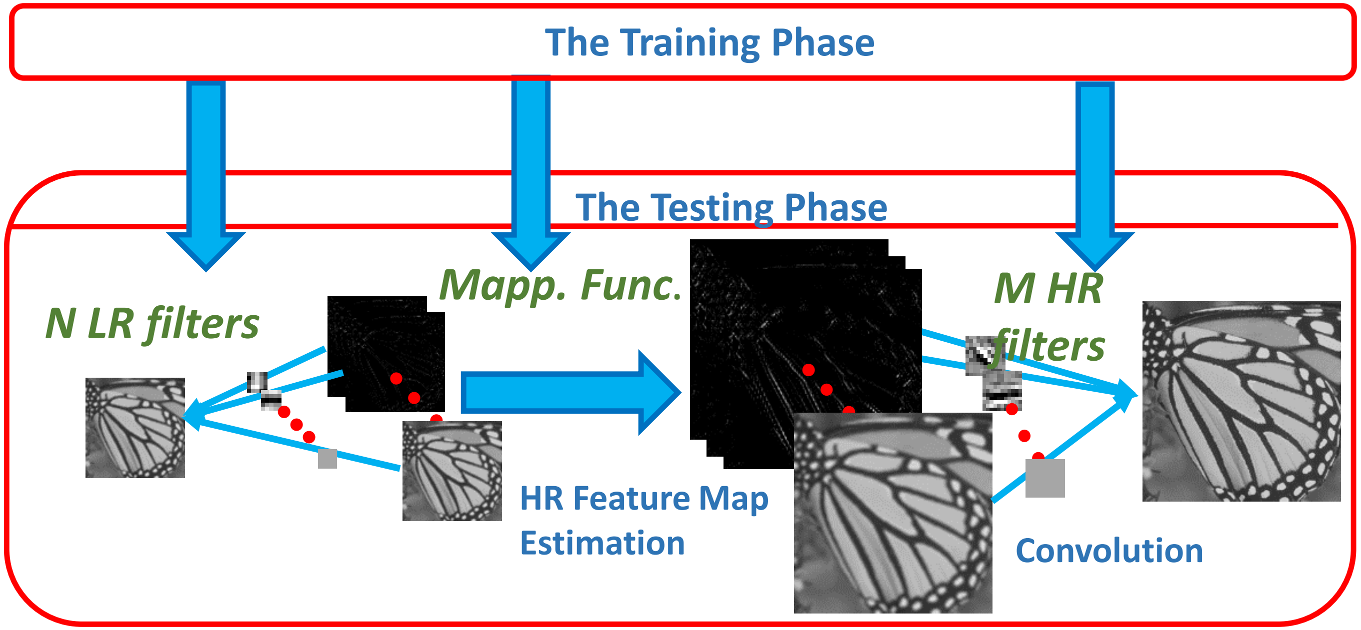


– Joint HR filter and mapping function learning

$$\{f^h, W\} = \min_{f,w} \| X - \sum_{j=1}^{M} f_j^h \otimes g(Z_:^l; w_j) \|_F^2$$

$$s.t. \; \| f_j^h \|_F^2 \leq e; \quad w_j \succeq 0, |w_j|_1 = 1$$

Convolutional sparse coding for image SR



Convolutional sparse coding for image SR

Optimization: SA-ADMM

$$\{\boldsymbol{W}\} = \arg\min_{\boldsymbol{W}} \sum_{k=1}^{K} \| \boldsymbol{X}_k - \sum_{j=1}^{M} \boldsymbol{f}_j^h \otimes g(\boldsymbol{Z}_{k,:}^l; \boldsymbol{w}_j) \|_F^2, \quad s.t. \;\; \boldsymbol{w}_j \succeq 0, |\boldsymbol{w}_j|_1 = 1.$$

Denote by $\tilde{\boldsymbol{Z}}_i^l$ the upsampling of LR feature map

$$\tilde{\boldsymbol{Z}}_{k,i}^l(x', y') = \begin{cases} \boldsymbol{Z}_{k,i}^l(x, y) & if \; mod(x', factor) = 0 \; and \; mod(y', factor) = 0 \\ 0 & otherwise \end{cases},$$

then we have

$$[vec(\boldsymbol{Z}_{k,1}^h), vec(\boldsymbol{Z}_{k,2}^h), \ldots, vec(\boldsymbol{Z}_{k,M}^h)] = [vec(\tilde{\boldsymbol{Z}}_{k,1}^l), vec(\tilde{\boldsymbol{Z}}_{k,2}^l), \ldots, vec(\tilde{\boldsymbol{Z}}_{k,N}^l)] * \boldsymbol{W},$$

The original problem can be write as:

$$\{\boldsymbol{W}\} = \sum_{k=1}^{K} \arg\min_{\boldsymbol{W}} \| vec(\boldsymbol{X}) - [\boldsymbol{F}_1^h, \ldots, \boldsymbol{F}_M^h] * \begin{bmatrix} [vec(\tilde{\boldsymbol{Z}}_{k,1}^l), \ldots, vec(\tilde{\boldsymbol{Z}}_{k,N}^l)] & & \\ & \cdots & \\ & & [vec(\tilde{\boldsymbol{Z}}_{k,1}^l), \ldots, vec(\tilde{\boldsymbol{Z}}_{k,N}^l)] \end{bmatrix} * vec(\boldsymbol{W}) \|_F^2$$

$$s.t. \;\; \boldsymbol{w}_j \succeq 0, |\boldsymbol{w}_j|_1 = 1.$$

Convolutional sparse coding for image SR

Optimization: SA-ADMM

$$\{\boldsymbol{W}\} = \sum_{k=1}^{K} \arg\min_{\boldsymbol{W}} \|vec(\boldsymbol{X}) - [\boldsymbol{F}_1^h, \ldots, \boldsymbol{F}_M^h] * \begin{bmatrix} [vec(\tilde{\boldsymbol{Z}}_{k,1}^l), \ldots, vec(\tilde{\boldsymbol{Z}}_{k,N}^l)] & & \\ & \cdots & \\ & & [vec(\tilde{\boldsymbol{Z}}_{k,1}^l), \ldots, vec(\tilde{\boldsymbol{Z}}_{k,N}^l)] \end{bmatrix} * vec(\boldsymbol{W})\|_F^2$$

$$s.t.\ \ \boldsymbol{w}_j \succeq 0, |\boldsymbol{w}_j|_1 = 1.$$

$$\{\boldsymbol{W}\} = \sum_{k=1}^{K} \arg\min_{\boldsymbol{W}} \|vec(\boldsymbol{X}) - \boldsymbol{A} * vec(\boldsymbol{W})\|_F^2 \quad s.t.\ \ \boldsymbol{w}_j \succeq 0, |\boldsymbol{w}_j|_1 = 1.$$

SA-ADMM

$$vec(\boldsymbol{W})_{t+1} = [L vec(\bar{\boldsymbol{W}})_t - \rho(\boldsymbol{T}_t - \boldsymbol{S}_t) - \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{A}_k^T(\boldsymbol{A}_k vec(\boldsymbol{W}_{\tau_j(t)}) - \boldsymbol{X}_k)]/(\rho + L)$$

$$\boldsymbol{S}_{t+1} = argmin_{\boldsymbol{S}} \frac{\rho}{2}\|\boldsymbol{W}_{t+1} + \boldsymbol{T}_t - \boldsymbol{S}\|^2, \quad s.t.\ \ \boldsymbol{s}_j \succeq 0, \sum \boldsymbol{s}_j = 1$$

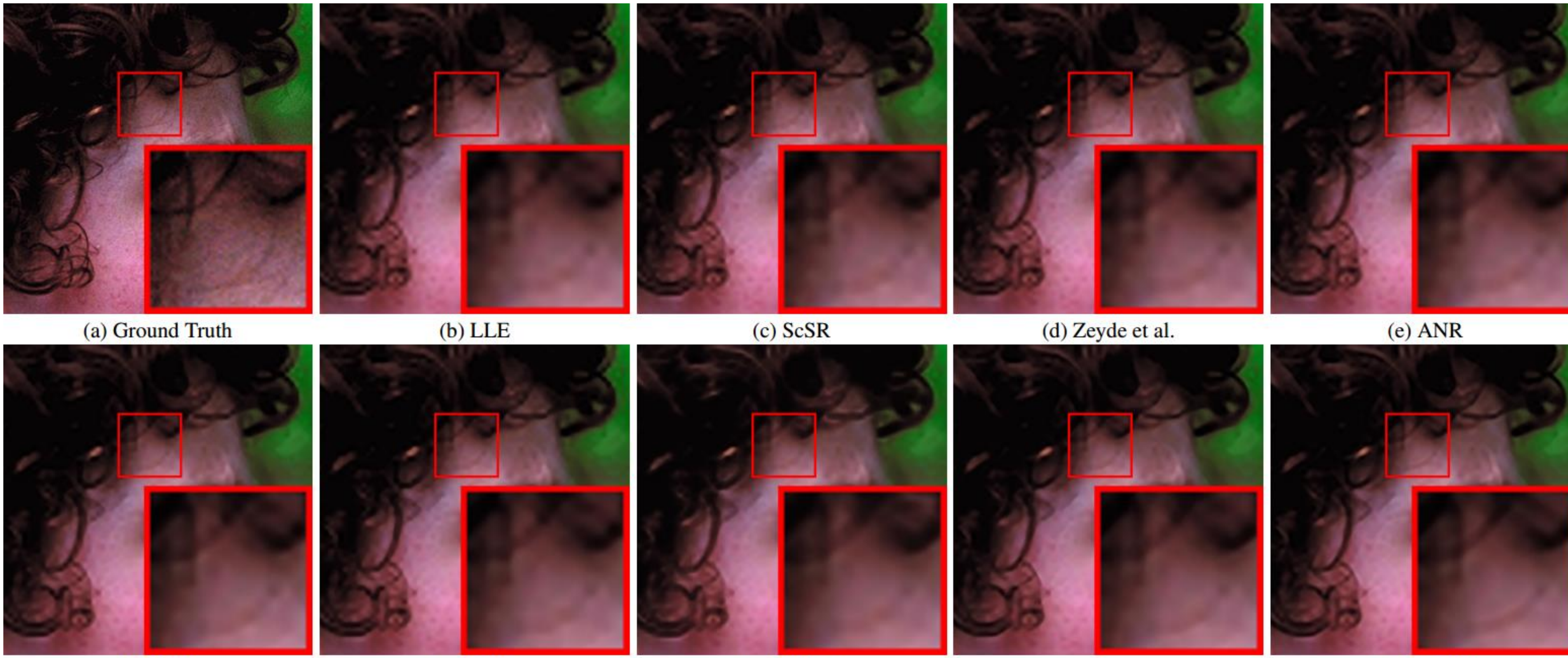
$$\boldsymbol{T}_{t+1} = \boldsymbol{T}_t + \boldsymbol{W}_{t+1} - \boldsymbol{S}_{t+1}$$

Convolutional sparse coding for image SR

(a) Ground Truth (b) LLE (c) ScSR (d) Zeyde et al. (e) ANR

(f) BPJDL (g) DPSR (h) CNN-SR (i) A+ (j) CSC-SR

Convolutional sparse coding for image SR



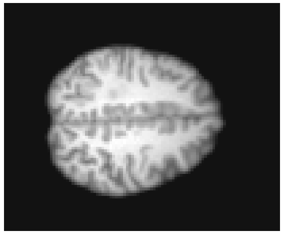
(a) Ground Truth (b) LLE (c) ScSR (d) Zeyde et al. (e) ANR

(f) BPJDL (g) DPSR (h) CNN-SR (i) A+ (j) CSC-SR

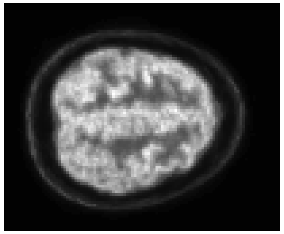
Guided Image Enhancement via Weighted Analysis Sparsity

Dependency Modeling

- Dependent image data



MRI



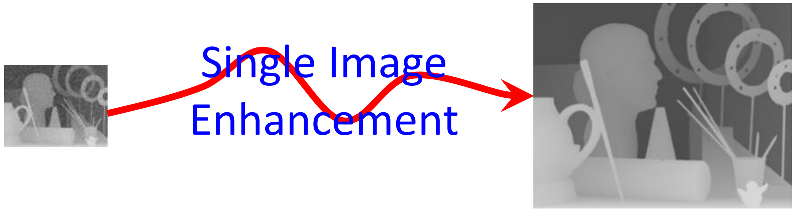
PET

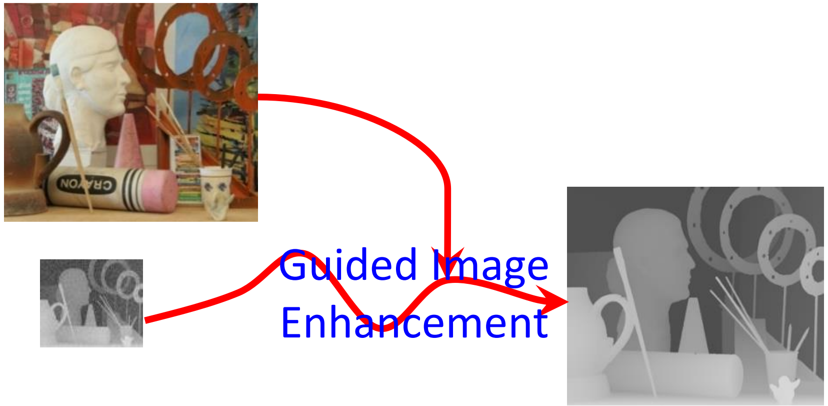
Depth

RGB

...

- Guided enhancement





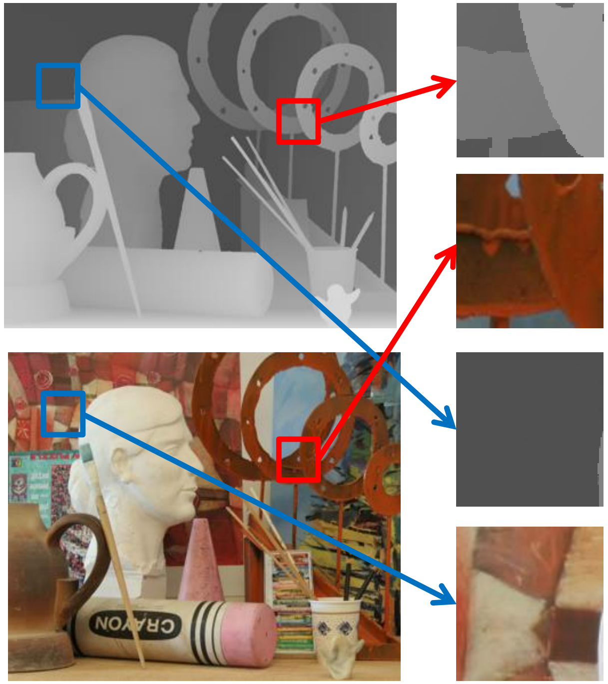
Dependency Modeling

- Previous Arts
  - 1st order method: co-different

$$\sum_i \sum_{j \in S(i)} (\boldsymbol{x}_i - \boldsymbol{x}_j)^2 \varphi(\boldsymbol{g}_i - \boldsymbol{g}_j)$$

$$\sum_i \sum_{j \in S(i)} (1 - \rho(\boldsymbol{x}_i - \boldsymbol{x}_j)) \varphi(\boldsymbol{g}_i - \boldsymbol{g}_j)$$

  - 2nd order method: TGV
  - Other priors: Non-local mean
  - Data-driven method: joint dictionary learning



Dependency Modeling

- Weighted Analysis Sparse Representation Model

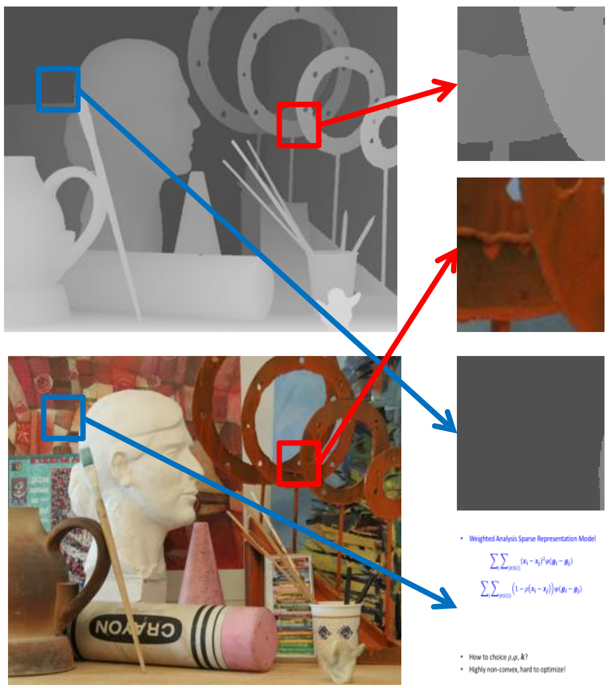
$$\sum_i \sum_{j \in S(i)} (\boldsymbol{x}_i - \boldsymbol{x}_j)^2 \varphi(\boldsymbol{g}_i - \boldsymbol{g}_j)$$

$$\sum_i \sum_{j \in S(i)} \left(1 - \rho(\boldsymbol{x}_i - \boldsymbol{x}_j)\right) \varphi(\boldsymbol{g}_i - \boldsymbol{g}_j)$$

Generalize the model: from one-order point –wise relationship to high-order local prior.

$$\hat{x} = argmin_X f(\boldsymbol{x}, \boldsymbol{y}) + \sum \rho(\boldsymbol{k}_x^i * \boldsymbol{x}) \odot \varphi(\boldsymbol{k}_g^i * \boldsymbol{g})$$

- How to choice ρ,φ, $\boldsymbol{k}$?
- Highly non-convex, hard to optimize!



Guided image enhancement via weighted analysis sparsity

- Task-driven training of stage-wise parameters
  - Solving weighted analysis sparse representation model with gradient descent, we have:

$$\hat{x} = argmin_x f(\boldsymbol{x}, \boldsymbol{y}) + \sum \rho(\boldsymbol{k}_x^i * \boldsymbol{x}) \odot \varphi(\boldsymbol{k}_g^i * \boldsymbol{g})$$

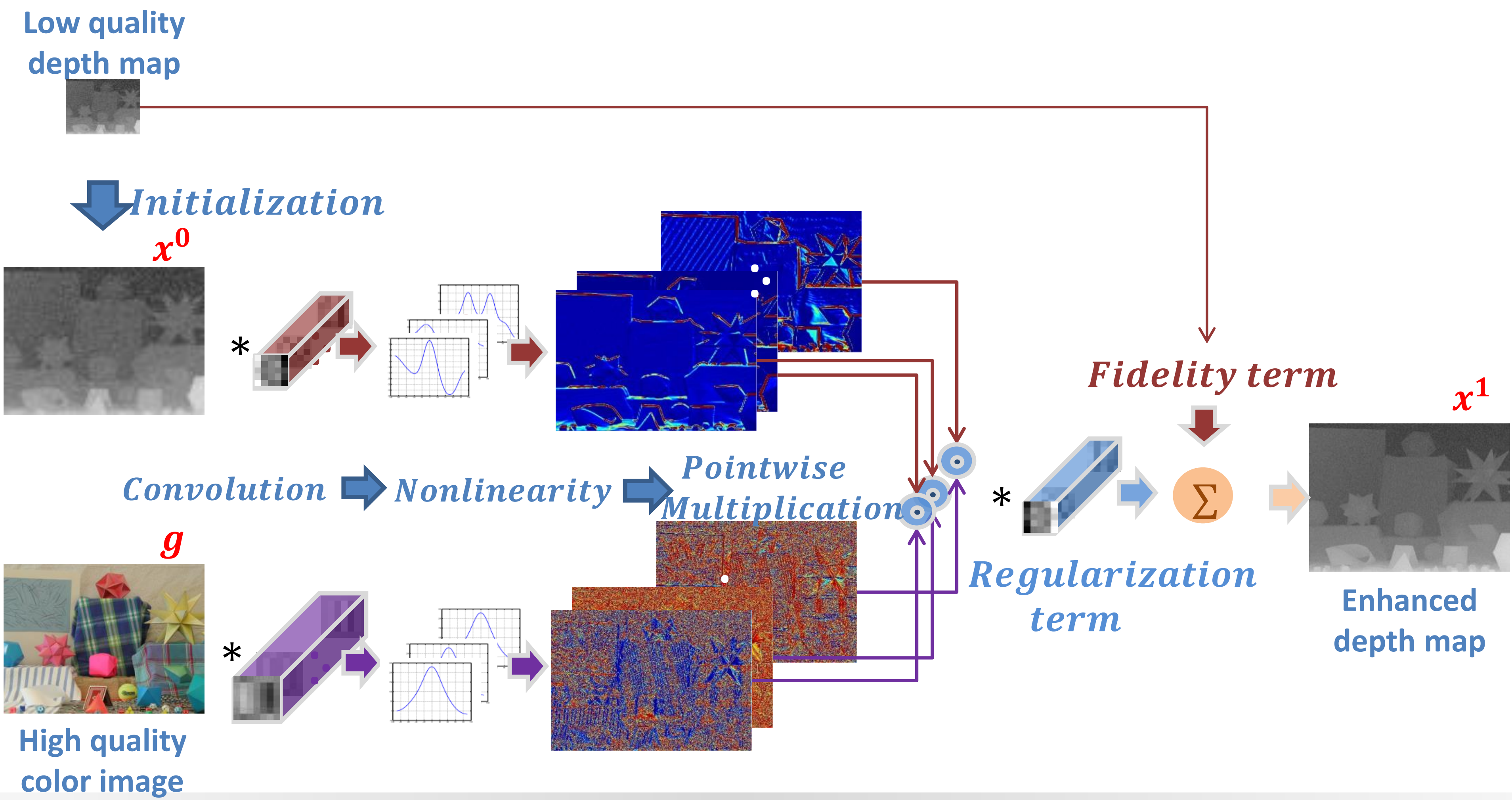
$$\boldsymbol{x}^{t+1} = \boldsymbol{x}^t - \tau \left(\Delta f(\boldsymbol{x}^t, \boldsymbol{y}) + \sum \boldsymbol{k}_x^{i\,T} \rho^{\boldsymbol{i}'} (\boldsymbol{k}_x^i * \boldsymbol{x}) \odot \varphi(\boldsymbol{k}_g^i * \boldsymbol{g}) \right)$$

  - Stage-wise parameter training

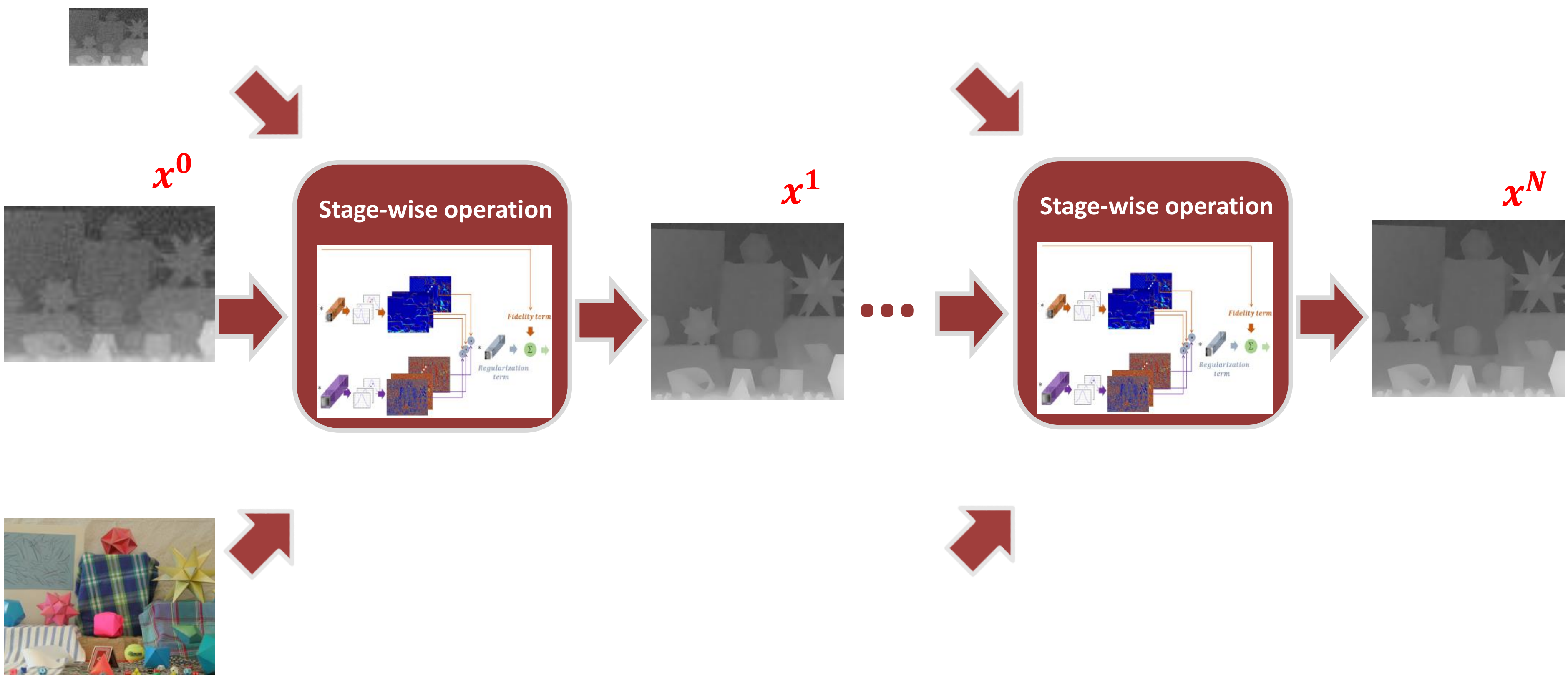
$$\boldsymbol{min}_{\boldsymbol{k}_x, \boldsymbol{k}_g, \rho, \varphi} \boldsymbol{loss}\left(\boldsymbol{x}^{t+1}\left(\boldsymbol{x}^t; \boldsymbol{k}_x, \boldsymbol{k}_g, \rho, \varphi\right) - \boldsymbol{x}^{gt}\right)$$

$$\boldsymbol{s.t.} \quad \boldsymbol{x}^{t+1}\left(\boldsymbol{x}^t; \boldsymbol{k}_x, \boldsymbol{k}_g, \rho, \varphi\right) = \boldsymbol{x}^t - \tau \left(\Delta f(\boldsymbol{x}^t, \boldsymbol{y}) + \sum \boldsymbol{k}_x^{i\,T} \rho^{\boldsymbol{i}'} (\boldsymbol{k}_x^i * \boldsymbol{x}) \odot \varphi(\boldsymbol{k}_g^i * \boldsymbol{g}) \right)$$

Guided image enhancement via weighted analysis sparsity

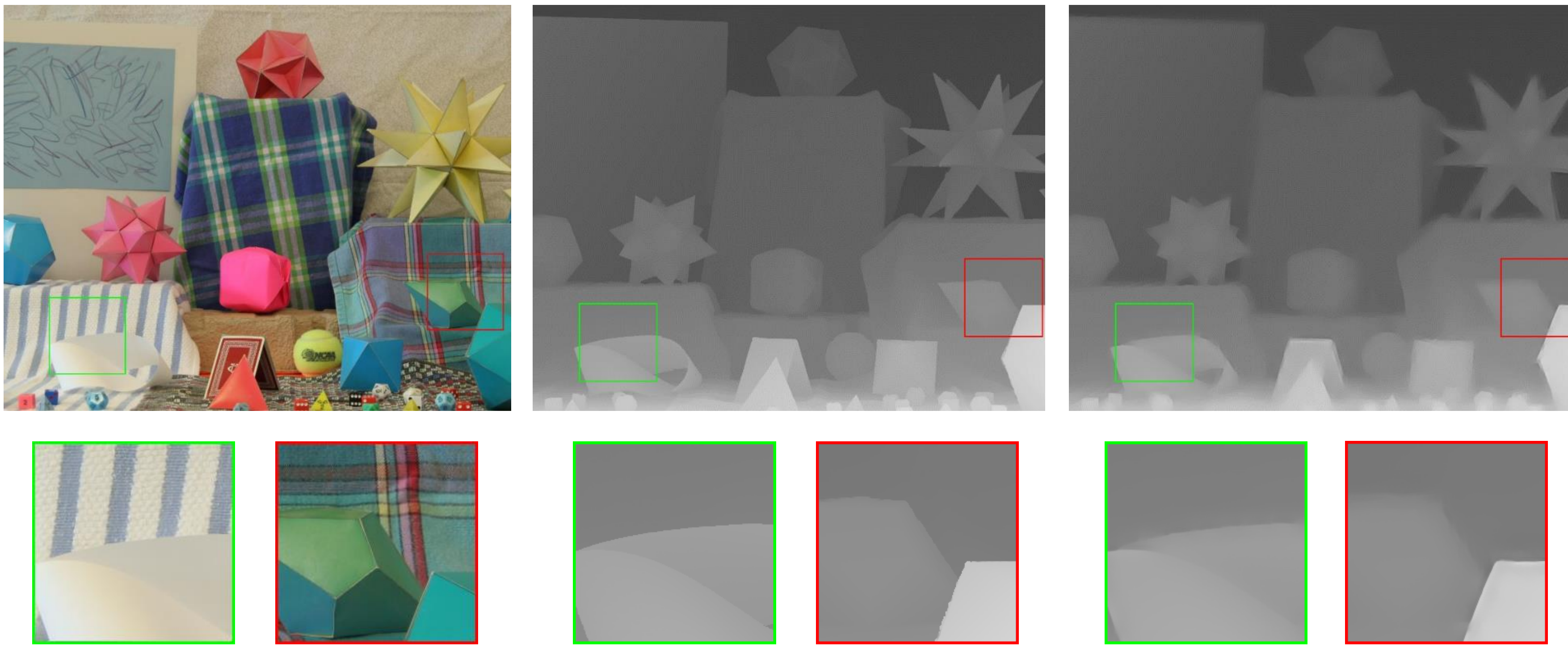


Guided image enhancement via weighted analysis sparsity



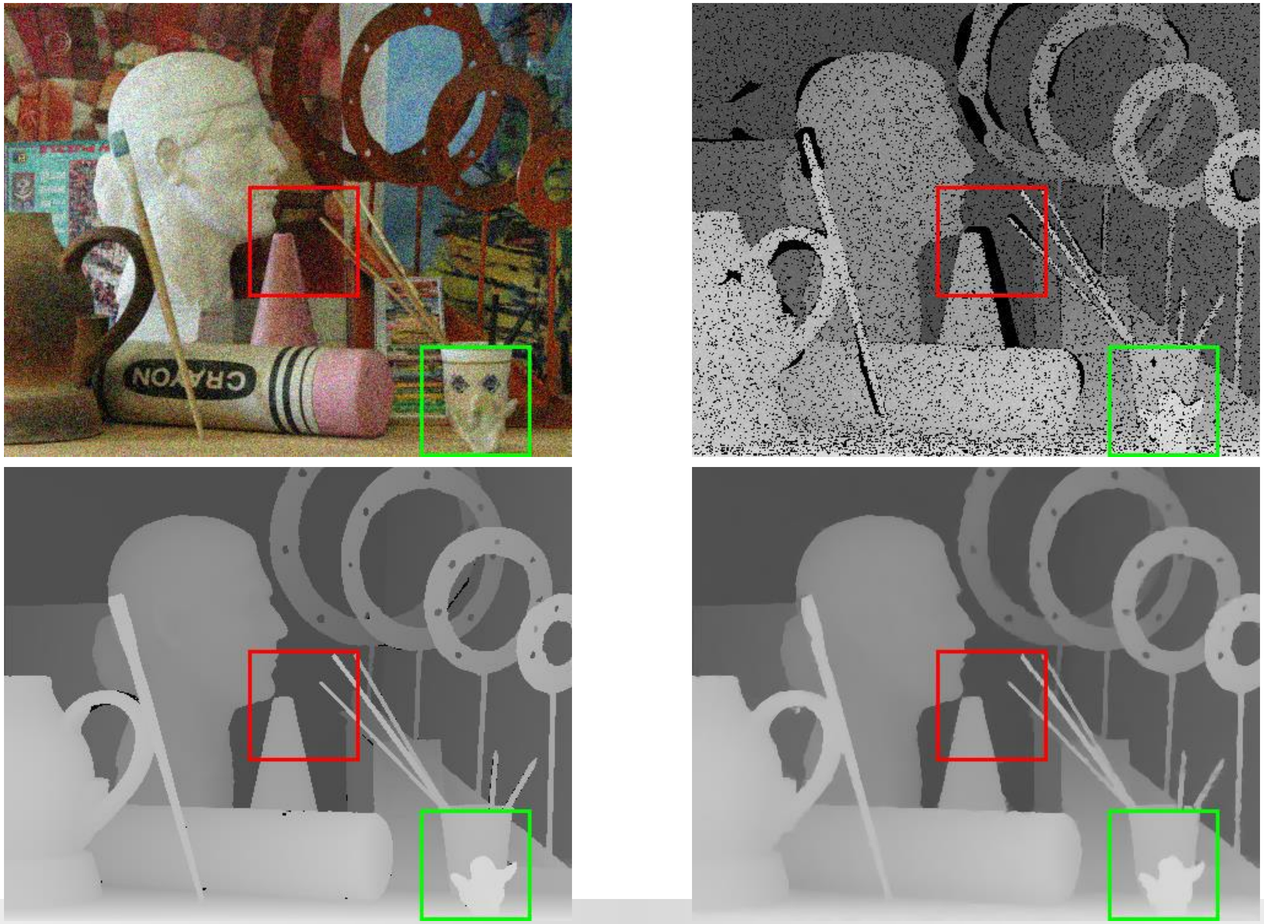
Guided Image Enhancement via Weighted Analysis Sparsity

- Experimental results



Guided Image Enhancement via Weighted Analysis Sparsity

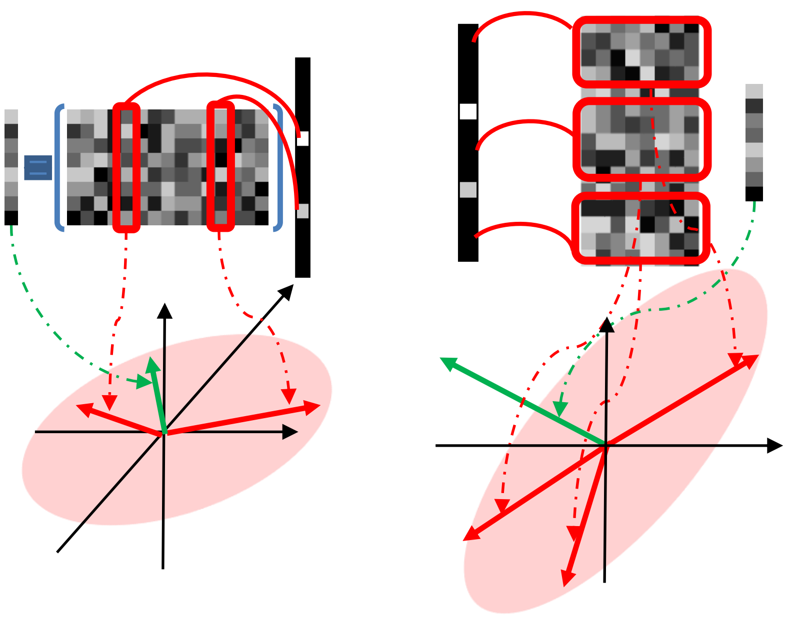
- Experimental results



Ongoing and Future Works

Image Separation without Training Data

- Complementary Property of ASR and SSR



- Layer Separation

$$min_{u,v} f(\boldsymbol{y}-\boldsymbol{u}-\boldsymbol{v})+\rho_S(\boldsymbol{u})+\rho_A(\boldsymbol{v})$$

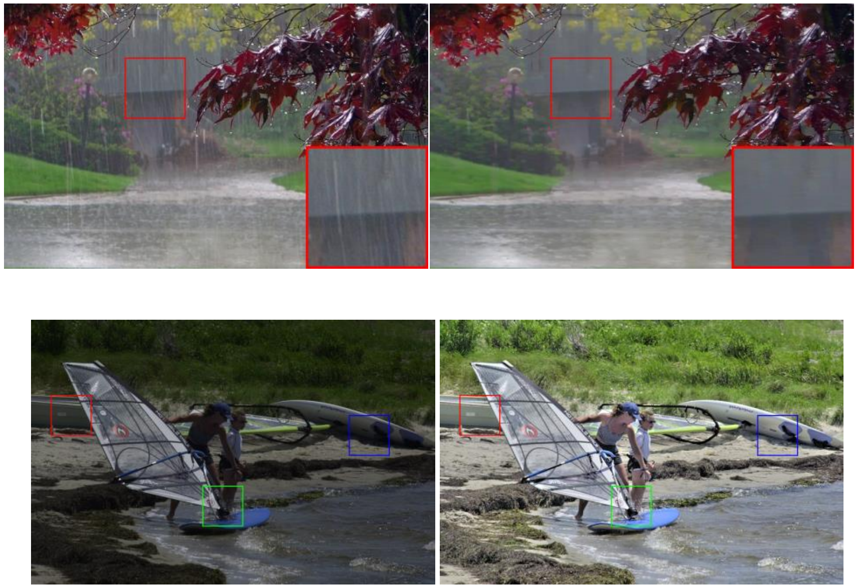


Image Restoration with Deep Denoisers

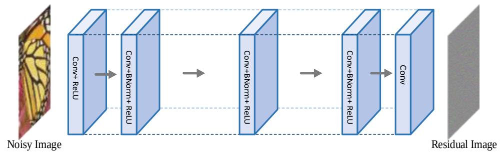
- Half Quadratic Splitting

$$min_x f(\boldsymbol{y}, \boldsymbol{x}) + \rho(\boldsymbol{x})$$

Half Quadratic Splitting

$$min_{x,s} f(\boldsymbol{y}, \boldsymbol{x}) + \lambda||\boldsymbol{x} - \boldsymbol{s}||_F^2 + \rho(\boldsymbol{s})$$

$$min_x f(\boldsymbol{y}, \boldsymbol{x}) + \lambda||\boldsymbol{x} - \boldsymbol{s}||_F^2 \qquad min_s \lambda||\boldsymbol{x} - \boldsymbol{s}||_F^2 + \rho(\boldsymbol{s})$$



Deep Denoiser

Optimization Inspired Network Structure Design

- State-of-the-art Performance Has been Achieved by Deep Models
  - Non-blind Super-Resolution
  - Gaussian Denoising
  - Non-blind Deblur
- More Complex Restoration problems
  - Blind Deblur, SR, Denoising?

Related Publications and References

Related Publication

- **S. Gu**, W. Zuo, Q. Xie, D. Meng, X. Feng, L. Zhang. "Convolutional Sparse Coding for Image Super-resolution," In **ICCV 2015**.
- **S. Gu**, W. Zuo, S. Guo, Y. Chen, C. Chen and L. Zhang, "Learning Dynamic Guidance for Depth Image Enhancement," To appear in **CVPR 2017**.
- K. Zhang, W. Zuo, **S. Gu** and L. Zhang, "Learning Deep CNN Denoiser Prior for Image Restoration," To appear in **CVPR 2017**.
- **S. Gu**, et. al. Joint Convolutional Analysis and Synthesis Sparse Representation for Single Image Layer Separation. Submitted.

Related Publications and References

References

- Elad, M., Milanfar, P., Rubinstein, R. Analysis versus synthesis in signal priors. Inverse problems 2007.
- Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. In CVPR 2005.
- M. Aharon, M. Elad, A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. TSP 2006.
- K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. TIP, 2007.
- J. Marial, F. Bach, J. P once, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In ICCV 2009.
- W. Dong, L. Zhang, and G. Shi, "Centralized Sparse Representation for Image Restoration," In *ICCV 2011*.
- Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena 1992.

Related Publications and References

References

- Zhu, S.C., Wu, Y., Mumford, D.: Filters, random fields and maximum entropy (frame): Towards a unified theory for texture modeling. IJCV 1998.
- Roth, S., Black, M.J.: Fields of experts. IJCV 2009.
- Schmidt, U., Roth, S.: Shrinkage fields for effective image restoration. In: CVPR. (2014)
- Chen, Y., Yu, W., Pock, T.: On learning optimized reaction diffusion processes for effective image restoration. In: CVPR. (2015)
- Rubinstein, Ron, Tomer Peleg, and Michael Elad. "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model." *TSP 2013*.
- Zoran D, Weiss Y. From learning models of natural image patches to whole image restoration. In: ICCV 2011.
- M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus. Deconvolutional networks. In CVPR, 2010.
- B. Wohlberg. Efficient convolutional sparse coding. In ICASSP, 2014.
- L. W. Zhong and J. T. Kwok. Fast stochastic alternating direction method of multipliers. In ICML, 2013.

Related Publications and References

References

- J. Diebel and S. Thrun. An application of markov random fields to range sensing. In NIPS, 2005.
- B. Ham, M. Cho, and J. Ponce. Robust image filtering using joint static and dynamic guidance. In CVPR, 2015.
- D. Ferstl, C. Reinbacher, R. Ranftl, M. Ruther, and H. Bischof. Image guided depth upsampling using anisotropic total generalized variation. In ICCV, 2013.
- J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon. High quality depth map upsampling for 3d-tof cameras. In ICCV, 2011.

THANKS!