Coupled Dictionary Learning for Multi-modal Image Super-resolution

Pingfan Song⋆ João Mota⋆ Nikos Deligiannis† Miguel Rodrigues⋆

⋆ University College London, UK
† Vrije Universiteit Brussel, Belgium

7th December 2016

UCL ENGINEERING
Change the world

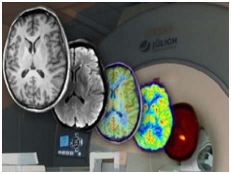
Introduction/ Motivation

- Multi-modal data available

  e.g., remote sensing, medical imaging, relic digitalization.

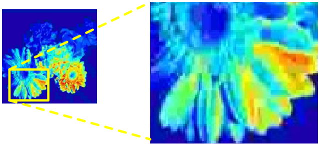
  exhibiting relationship, such as similarity and complementarity

Remote Sensing



Medical Imaging

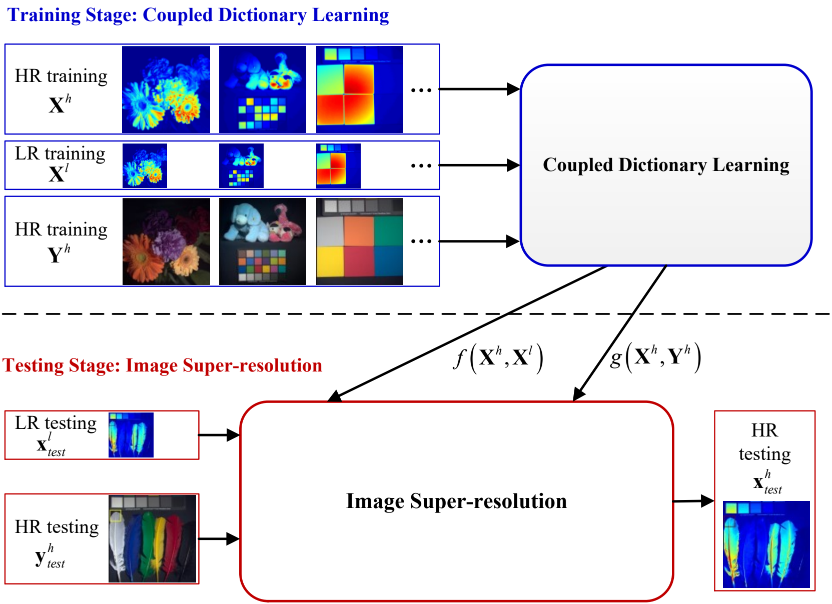
Relic Digitalization



Modality of low resolution

Modality of high resolution

Overview: Multi-modality Image Super-resolution



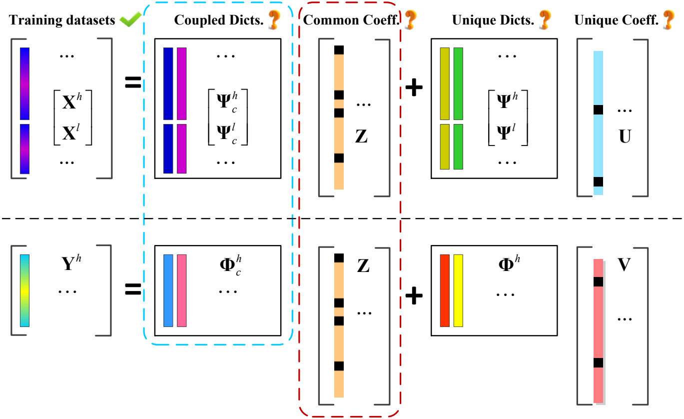
Training Stage: Coupled Dictionary Learning

Multi-modal Data Model:

$$\mathbf{X}^h = \Psi_c^h \mathbf{Z} + \Psi^h \mathbf{U} \quad (1)$$

$$\mathbf{X}^l = \Psi_c^l \mathbf{Z} + \Psi^l \mathbf{U} \quad (2)$$

$$\mathbf{Y}^h = \Phi_c^h \mathbf{Z} + \Phi^h \mathbf{V} \quad (3)$$



Training Stage: Coupled Dictionary Learning

$$\underset{\substack{\Psi_c,\Psi,\Phi_c,\Phi\\ \mathbf{Z},\mathbf{U},\mathbf{V}}}{\text{minimize}} \quad f = \sum_i \left\| \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} - \begin{bmatrix} \Psi_c & \Psi & \mathbf{0} \\ \Phi_c & \mathbf{0} & \Phi \end{bmatrix} \begin{bmatrix} \mathbf{z}_i \\ \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} \mathbf{z}_i \\ \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \right\|_1 \tag{4}$$

- (4) is NON-CONVEX for both $\{\Psi_c, \Psi, \Phi_c, \Phi\}$ and $\{\mathbf{Z}, \mathbf{U}, \mathbf{V}\}$.
- Alternating minimization between

  sparse coding: $\underset{\mathbf{Z},\mathbf{U},\mathbf{V}}{\text{minimize}} \quad f$

  and

  dictionary update: $\underset{\Psi_c,\Psi,\Phi_c,\Phi}{\text{minimize}} \quad f$

- Learn a dictionary with particular structure requirements.

Training Stage: Coupled Dictionary Learning

$$\underset{\substack{\Psi_c,\Psi,\Phi_c,\Phi\\ \mathbf{Z},\mathbf{U},\mathbf{V}}}{\text{minimize}} \quad f = \sum_i \left\| \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \end{bmatrix} - \begin{bmatrix} \Psi_c & \Psi & \mathbf{0} \\ \Phi_c & \mathbf{0} & \Phi \end{bmatrix} \begin{bmatrix} \mathbf{z}_i \\ \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} \mathbf{z}_i \\ \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \right\|_1 \qquad (4)$$

- (4) is NON-CONVEX for both $\{\Psi_c, \Psi, \Phi_c, \Phi\}$ and $\{\mathbf{Z}, \mathbf{U}, \mathbf{V}\}$.
- Alternating minimization between

  sparse coding: $\underset{\mathbf{Z},\mathbf{U},\mathbf{V}}{\text{minimize}} \quad f$

  and

  dictionary update: $\underset{\Psi_c,\Psi,\Phi_c,\Phi}{\text{minimize}} \quad f$

- Learn a dictionary with particular structure requirements.

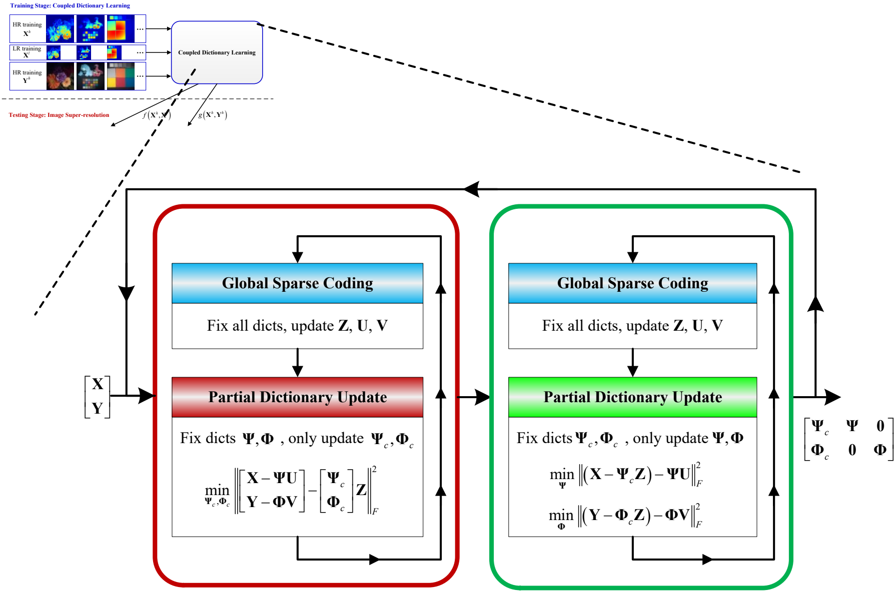
- dictionary update: split the whole problem into three parts so that the whole dictionary is updated part by part.

$$\min_{\Psi_c,\Phi_c} \quad \left\| \begin{bmatrix} \mathbf{X} - \Psi\mathbf{U} \\ \mathbf{Y} - \Phi\mathbf{V} \end{bmatrix} - \begin{bmatrix} \Psi_c \\ \Phi_c \end{bmatrix} \mathbf{Z} \right\|_F^2$$

$$\min_{\Psi} \quad \|(\mathbf{X} - \Psi_c\mathbf{Z}) - \Psi\mathbf{U}\|_F^2$$

$$\min_{\Phi} \quad \|(\mathbf{Y} - \Phi_c\mathbf{Z}) - \Phi\mathbf{V}\|_F^2$$

Training Stage: Coupled Dictionary Learning



Testing Stage: Image Super-resolution

Training Stage: Coupled Dictionary Learning

$$f\left(\mathbf{X}^h,\mathbf{X}^l\right)=\left\{\boldsymbol{\Psi}^h_c,\boldsymbol{\Psi}^h,\boldsymbol{\Psi}^l_c,\boldsymbol{\Psi}^l\right\}$$

$$g\left(\mathbf{X}^h,\mathbf{Y}^h\right)=\left\{\boldsymbol{\Psi}^h_c,\boldsymbol{\Psi}^h,\boldsymbol{\Phi}^h_c,\boldsymbol{\Phi}^h\right\}$$

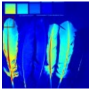
Testing Stage: Image Super-resolution

LR testing $\mathbf{x}^l_{test}$

HR testing $\mathbf{y}^h_{test}$

Image Super-resolution

HR testing $\mathbf{x}^h_{test}$



Sparse Coding

$$\min_{\mathbf{z},\mathbf{u},\mathbf{v}} \|\mathbf{z}\|_0+\|\mathbf{u}\|_0+\|\mathbf{v}\|_0$$

$$\text{s.t. } \mathbf{x}^l_{test}=\boldsymbol{\Psi}^l_c\mathbf{z}+\boldsymbol{\Psi}^l\mathbf{u}$$

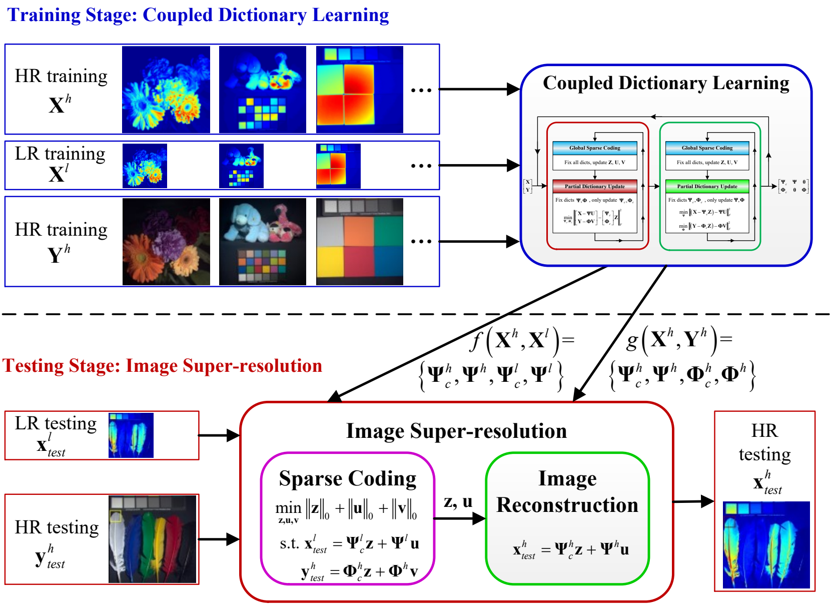
$$\mathbf{y}^h_{test}=\boldsymbol{\Phi}^h_c\mathbf{z}+\boldsymbol{\Phi}^h\mathbf{v}$$

z, u

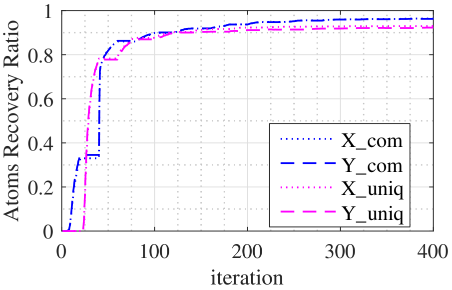
Image Reconstruction

$$\mathbf{x}^h_{test}=\boldsymbol{\Psi}^h_c\mathbf{z}+\boldsymbol{\Psi}^h\mathbf{u}$$

Summary for Multi-modality Image Super-resolution

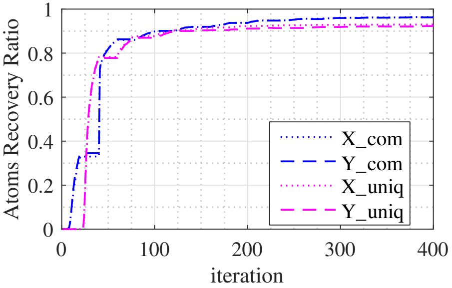


Experiments/ Simulation Results

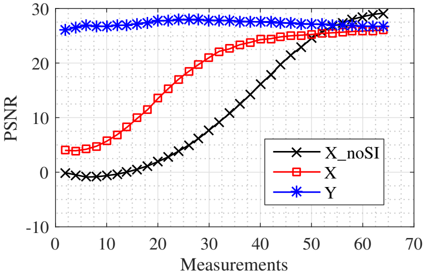


Coupled Dictionary Learning Simulation

Experiments/ Simulation Results

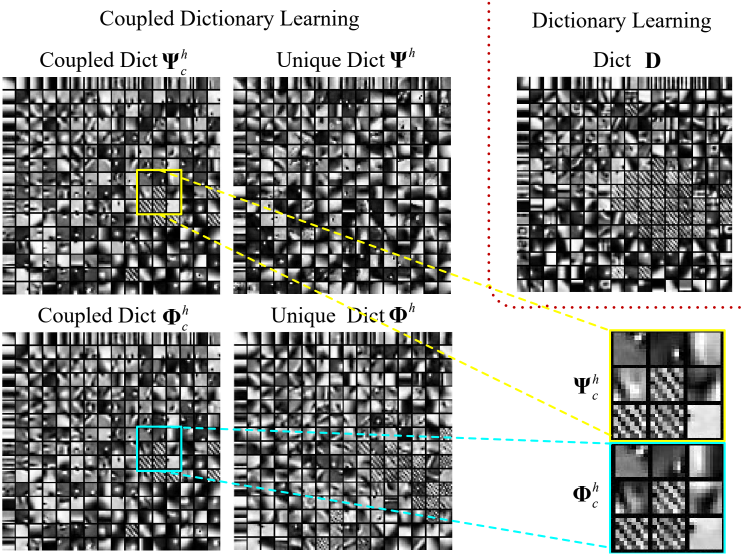


Coupled Dictionary Learning Simulation

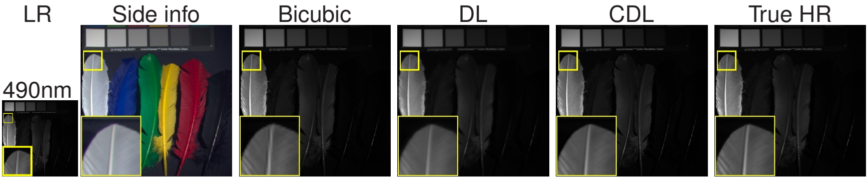


Super-resolution Simulation

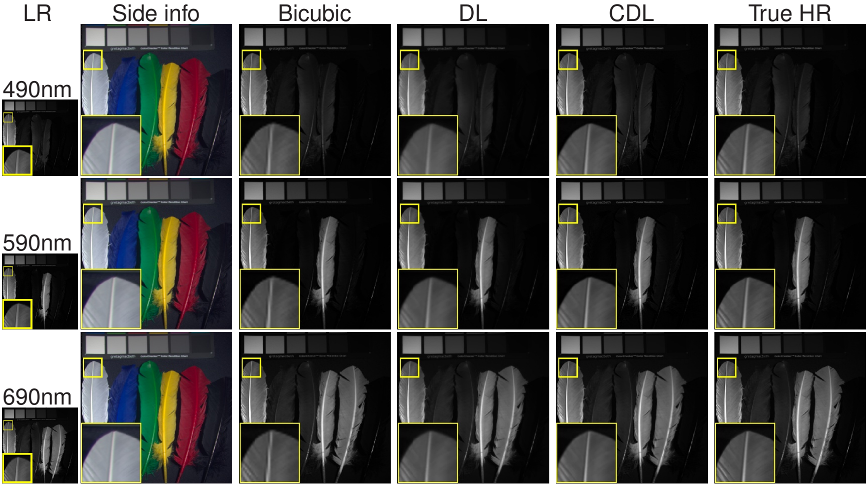
Experiments/ Practical CDL and DL results



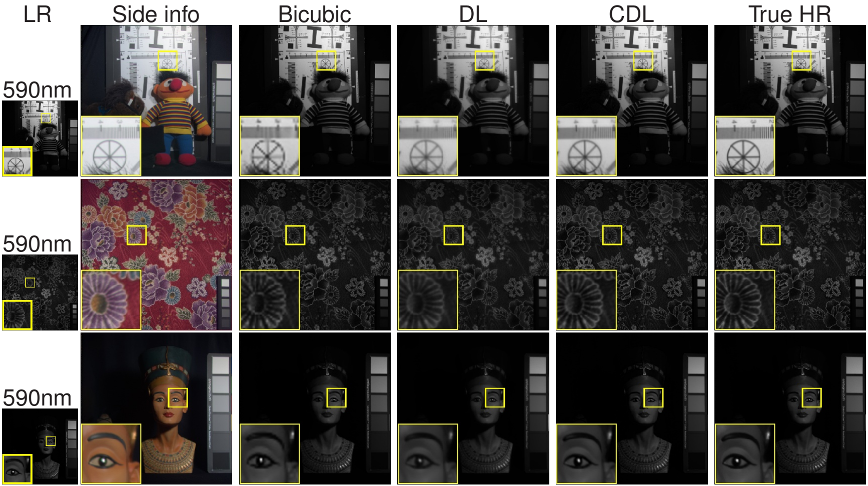
Experiments/ Practical super-resolution results



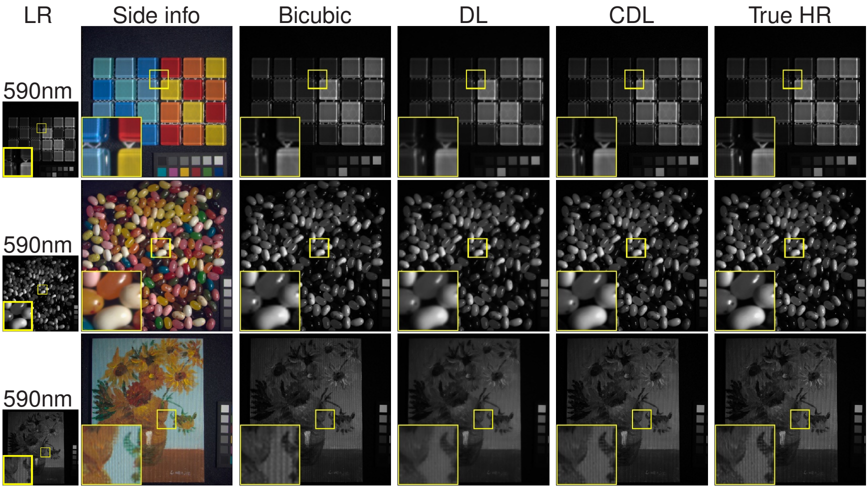
Experiments/ Practical super-resolution results



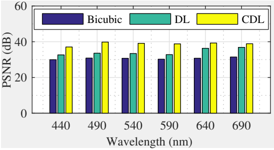
Experiments/ More results



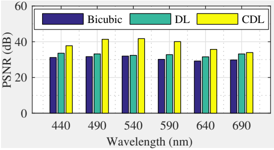
Experiments/ More results



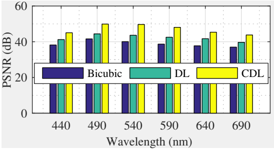
Experiments/ Practical super-resolution results



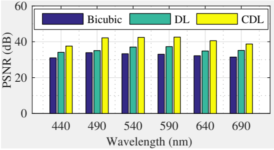
(a) Img No.1



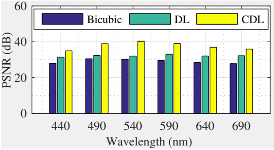
(b) Img No.2



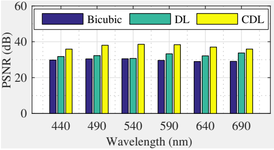
(c) Img No.3



(d) Img No.4



(e) Img No.5



(f) Img No.6

Figure 1：SR performance in terms of PSNR

Time spent in dictionary training phase (64 x 256 size):

DL — 3.82 min, CDL — 34.88 min.

Time spent in image super-resolution phase (one 512 x 512 image):

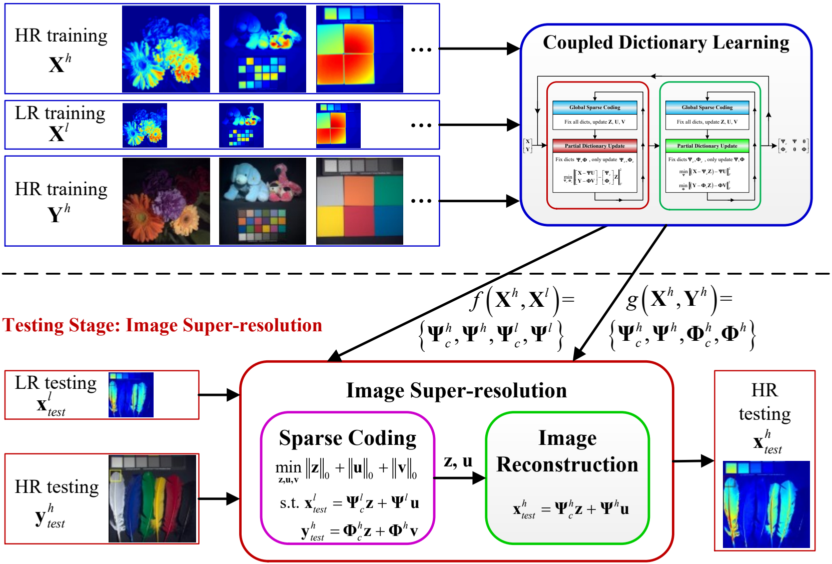
bicubic — 0.0035s, DL based SR — 59.89s, CDL based SR — 259.5s.

Summary and Conclusion

- Scenarios where multi-modal data is available
  e.g., remote sensing, medical imaging, relic digitalization.

- Multi-modal image super-resolution framework
  - Training stage:
    Multi-modal data model, Coupled dictionary learning (CDL)
    capture the inter-relationship among modalities.
  - Testing stage:
    Image super-resolution
    exploit the learned mapping relationship to enhance the resolution of LR images with corresponding HR modality as side information.

Training Stage: Coupled Dictionary Learning



Summary and Conclusion

- Scenarios where multi-modal data is available
  e.g., remote sensing, medical imaging, relic digitalization.

- Multi-modal image super-resolution framework
  - Training stage:
    Multi-modal data model, Coupled dictionary learning (CDL)
    capture the inter-relationship among modalities.
  - Testing stage:
    Image super-resolution
    exploit the learned mapping relationship to enhance the resolution of LR images with corresponding HR modality as side information.

- Outperform Bicubic interpolation and DL-based image SR

- Shortcoming: highly non-convex, time consuming, more computing resources.

Thanks!