Mathematical Principles Behind Fast Unsupervised Image Super-resolution

Chia-Hsiang Lin
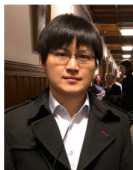
Intelligent Hyperspectral Computing Laboratory
Department of Electrical Engineering
National Cheng Kung University

E-mail: `chiahsiang.steven.lin@gmail.com`
Web: `https://sites.google.com/view/chiahsianglin/`

Presented at Department of Statistics, NCKU, Taiwan, September 10, 2020.

- **I**ntelligent **H**yperspectral **C**omputing **L**aboratory *(No magic, only basic!)*

Dr. Chia-Hsiang Lin (林家祥)

Assistant Professor, <u>Department of Electrical Engineering</u>

Assistant Professor, <u>Institute of Computer and Communication Engineering</u>

Assistant Professor, <u>Miin Wu School of Computing</u>

Advisor, <u>Intelligent Hyperspectral Computing Laboratory</u>

National Cheng Kung University

Honors and Awards

- **Top Performance Award**, Social Media Prediction Challenge, ACM Multimedia, 2020. ←
- **Outstanding Paper Award** from IPPR Conference on Computer Vision, Graphics, and Image Processing (CVGIP), 2020.
- **Prize Paper Award** (IGARSS'19) from <u>IEEE Geoscience & Remote Sensing Society</u>, 2020. ← **2500**
- 3rd Place, AIM Super-Resolution Challenge, IEEE International Conference on Computer Vision (ICCV), 2019. ← **8000**
- **Einstein Grant Award** from <u>Ministry of Science and Technology</u>, from 2018 to 2023.
- **Best Doctoral Dissertation Award** from <u>IEEE Geoscience & Remote Sensing Society</u>, 2016.
- **Outstanding Doctoral Dissertation Award** from <u>Chinese Image Processing and Pattern Recognition (IPPR) Society</u>, 2016.
- Honorary Member of <u>The Phi Tau Phi Scholastic Honor Society</u>, 2016 (for top 1% of the graduates of NTHU).
- International Visiting Scholarship, National Institutes of Health (NIH), MD, USA, 2015 (for visiting Virginia Tech).
- International Visiting Scholarship, NTHU, 2014 (for visiting CUHK).
- President Scholarship, NTHU, 2011 (for the best student of ICE, NTHU).
- One of the Top 5 Members of University's Go Team, NTHU, for 2009 and 2010 (清華圍棋校隊).
- Marathon Competition (completed 5 international full (42.195km) marathons) (全程馬拉松).
- Entrance Scholarship of Department of Electrical Engineering (EE), NTHU, 2005 (for top 10 best students of EE, NTHU).
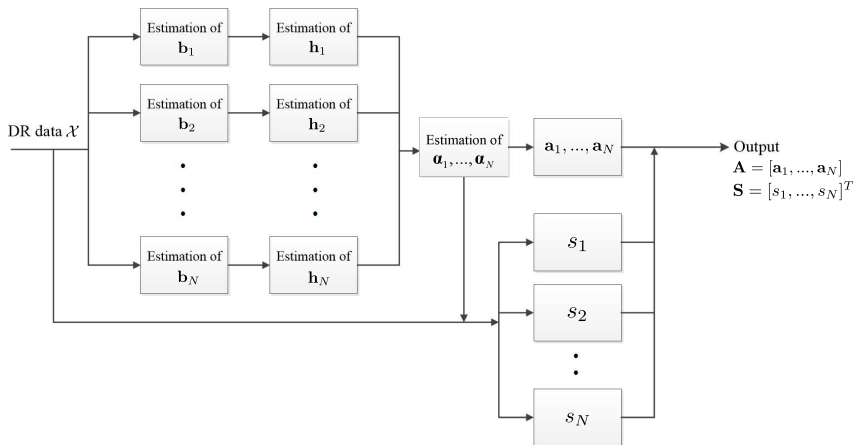- Primary Selected Student for the Physics Olympics National Team, 2004.

- Low Rank Model: $\mathbf{Z} = \mathbf{AS}$, or columnwisely $\mathbf{z}[n] = \mathbf{As}[n]$ [Lin'16a].
    - $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_N]$ is the spectral signature matrix.
    - $\mathbf{S}$ is the material abundance/distribution matrix.
    - The rank N, i.e., model order?
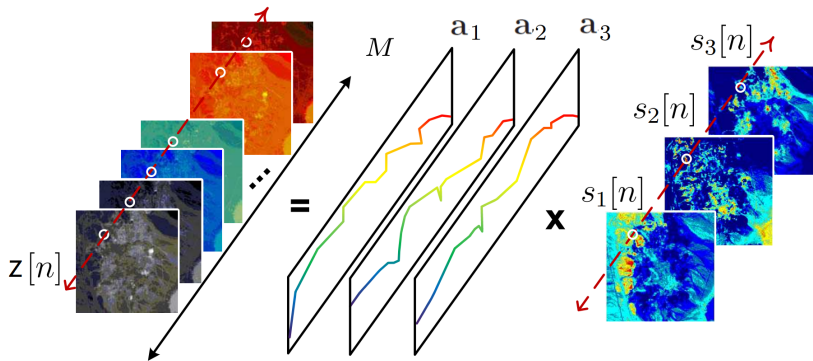
[Lin'16a] Chia-Hsiang Lin et al., "A fast hyperplane-based minimum-volume enclosing simplex algorithm for blind hyperspectral unmixing," *IEEE Trans. Signal Processing*, vol. 64, no. 8, pp. 1946-1961, Apr. 2016.

Even **without** *parallel computing, HyperCSI has been about 10,000 times faster than the benchmark MVC-NMF!*

Model Order Selection (MOS) *(Basic)*

- *A fundamental topic in Machine Learning!*
- Information-theoretic minimum description/code length (MDL).
  - The code length $\mathrm{MDL}(N) = -\log(f(\mathbf{X} \mid \Theta_{\mathsf{ML}}^{(N)})) + \frac{1}{2}D(\Theta^{(N)})\log(L)$ induces *very complicated* optimization (non-convex/non-smooth).
  - *Monte Carlo algorithm:* high-dimensional Gaussian-Dirichlet convoluted integral.
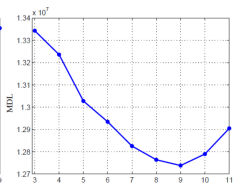- Effective for many research domains:
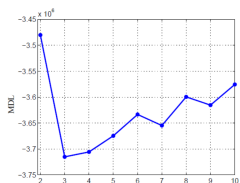
rat cell gene expression data
(3 cell types)
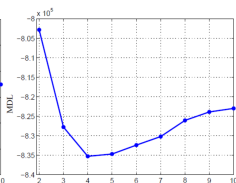
human blood microarray data
(4 immune origins)
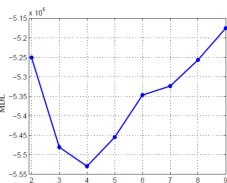
brain disease molecular data
(4 mRNAs)
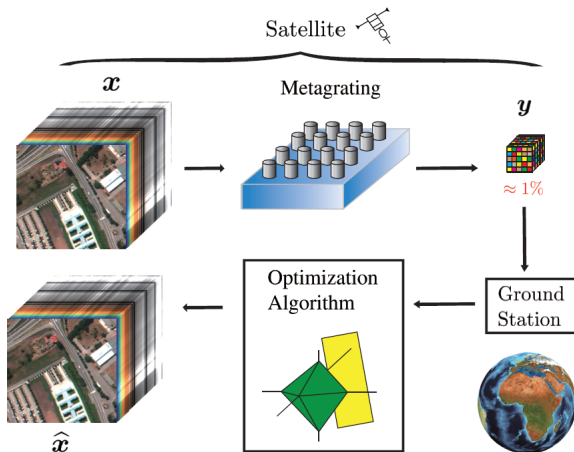
hyperspectral remote sensing data
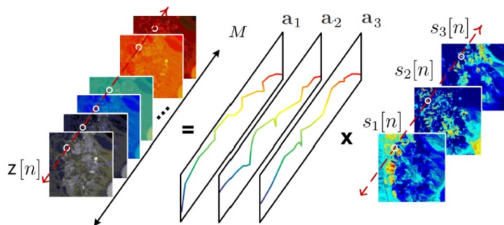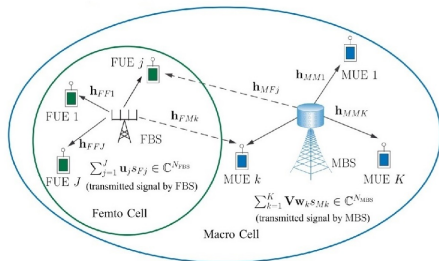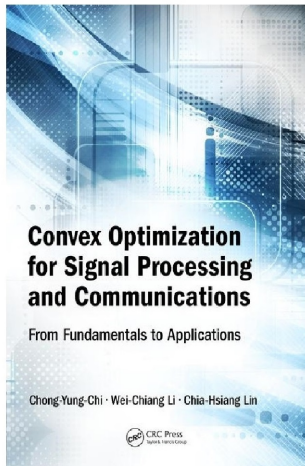(9 minerals)

(source codes available)

[Lin'17a] Chia-Hsiang Lin et al., "Detection of sources in non-negative blind source separation by minimum description length criterion," *IEEE Trans. Neural Networks and Learning Systems*, 2018.

- "All-addition hyperspectral compressed sensing for miniaturized satellite"
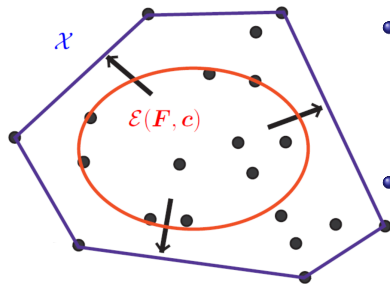
densely populated stripes

serious stripes
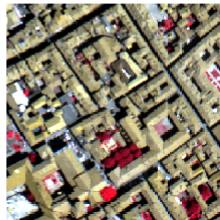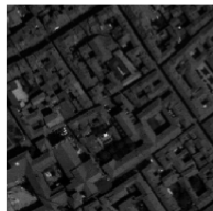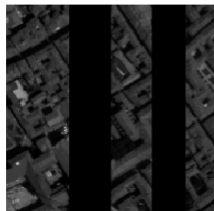
Löwner-John Ellipsoid (LJE) *(Basic)*

- It is the maximum-volume ellipsoid inscribed in the data-constructed convex hull.

  - $\mathcal{X}$: convex hull of dimension-reduced data.
  - $\mathcal{E}(\boldsymbol{F}, \boldsymbol{c}) \triangleq \{\boldsymbol{F}\boldsymbol{u} + \boldsymbol{c} \mid \|\boldsymbol{u}\|_2 \leq 1\}$, where $\boldsymbol{F}$ characterizes the semi-axis directions/lengths, and $\boldsymbol{c}$ is the ellipsoid center.
  - Formulation:

$$\max_{\boldsymbol{F} \in \mathbb{R}^{(N-1) \times (N-1)}, \boldsymbol{c} \in \mathbb{R}^{N-1}} \quad \text{vol}(\mathcal{E}(\boldsymbol{F}, \boldsymbol{c}))$$
$$\text{s.t.} \quad \mathcal{E}(\boldsymbol{F}, \boldsymbol{c}) \subseteq \mathcal{X}, \tag{1}$$

  - Solution:
    - $\mathcal{E}(\boldsymbol{F}, \boldsymbol{c}) = \mathcal{E}(\sqrt{\boldsymbol{F}\boldsymbol{F}^T}, \boldsymbol{c})$, so we can assume w.l.o.g. that $\boldsymbol{F} \in \mathbb{S}_{++}^{N-1}$.
    - $\text{vol}(\mathcal{E}(\boldsymbol{F}, \boldsymbol{c})) \propto \log \det(\boldsymbol{F})$, whose negative is convex on $\mathbb{S}_{++}^{N-1}$.
    - $\mathcal{E}(\boldsymbol{F}, \boldsymbol{c}) \subseteq \mathcal{X}$ is equivalent to a set of convex second-order cone constraints $\|\boldsymbol{F}\boldsymbol{b}_i\| \leq h_i - \boldsymbol{b}_i^T \boldsymbol{c}, \ \forall i = 1, \ldots, m$.

$\mathcal{X}$

$\mathcal{E}(\boldsymbol{F}, \boldsymbol{c})$
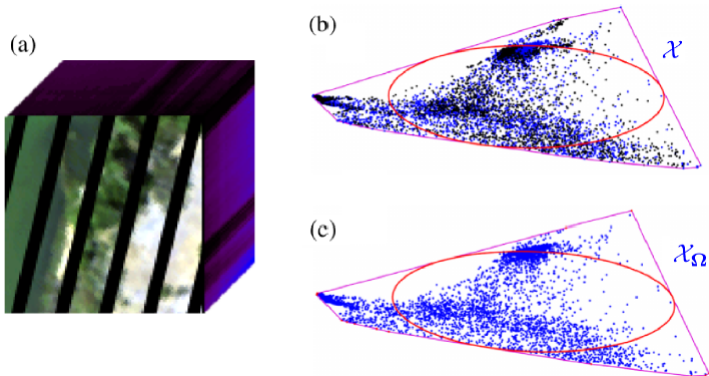
Figure 1: Topology mapping between complete and missing matrices.
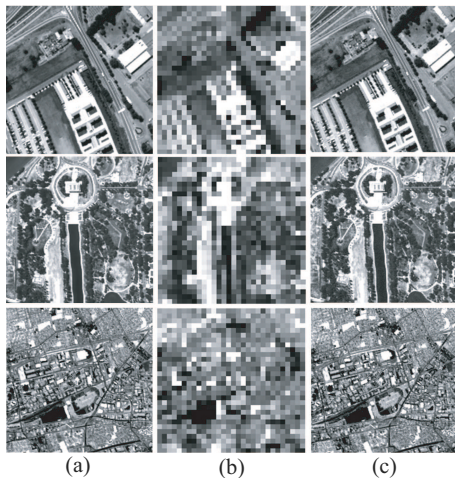
Figure: The 50th band of (a) $\mathbf{Z}$, (b) $\mathbf{Y}_h$ and (c) the inversed $\widehat{\mathbf{Z}}$ by CO-CNMF on 3 data: Pavia University (top), Washington DC (middle) and Moffett Field (bottom).

Data Fusion Theory *(Basic)*

- Hyperspectral data $\mathbf{Y}_h$:
  - high spectral resolution;
  - low spatial resolution;
  - bands: $M \gg M_m$
- Multispectral data $\mathbf{Y}_m$:
  - low spectral resolution;
  - high spatial resolution;
  - pixels: $L \gg L_h$
- Super-resolved data $\widehat{\mathbf{Z}}$:
  - high spectral resolution;
  - high spatial resolution.
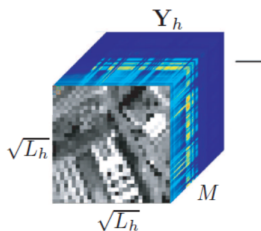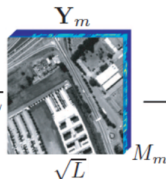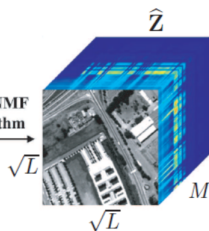
- Annual Data Fusion Contest, IEEE Geoscience and Remote Sensing Society (GRSS). *(Very important topic for remote sensing!)*
- State-of-the-art results in our recent paper [Lin'18a]! (source codes available)

[Lin'18a] Chia-Hsiang Lin et al., "A convex optimization based coupled non-negative matrix factorization algorithm for hyperspectral and multispectral data fusion," *IEEE Trans. Geoscience and Remote Sensing*, 2018.

超穎材料，讓你也能有件隱形斗篷！

2019/02/15　　　樂丕綱

超穎物質（metamaterials）是一種具有人造次波長結構的特殊材料，藉由介質參數的安排，任意操控光的傳播行為。因此飛機只要包裹在超穎材質中，光就會沿著超穎物質表面傳播，飛機前方的探測器就無法得知它的存在，此便如同隱形斗篷一般。超穎材質是當今重要的光電物理與奈米光學領域，本文將從材料折射談起，說明如何藉由介質係數控制來達到隱形效果。

There is no magic. There is only knowledge, more or less hidden.

Gene Wolfe

- Next, we focus on ESA's Sentinel-2 Satellite super-resolution problem.

Super-resolution in Remote Sensing

- **Why?** "Remote" v.s. "Spatial resolution" (classification).

Pavia University, Italy (ROSIS Sensor)

Super-resolved Image by CO-CNMF

- **How?** (sounds like magic)

  Direct acquisition? (too expensive, hardware limitation)

  Image fusion? (no high-resolution counterpart)

  $\implies$ *"Single image super-resolution (SISR)"*

Single Image Super-resolution (SISR)

- SISR for RGB images — a key topic for computer vision (CV) [Glasner'09]!
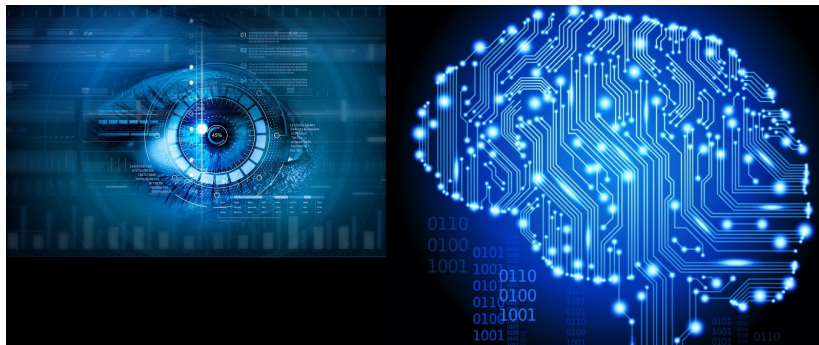
- SISR for Sentinel-2 (a new satellite) multi-spectral images is challenging due to
  - insufficient data to train the deep neural network;
  - multi-resolution images (see next page).

[Glasner'09] D. Glasner et al., "Super-resolution from a single image," *IEEE International Conference on Computer Vision (ICCV)*, 2009, pp. 349-356.

Sentinel-2 Satellite

- Launched by European Space Agency (ESA) with the following specifications:

Band	B1	B2	B3	B4	B5	B6	B7	B8	B8a	B9	B10	B11	B12
Central Wavelength (nm)	443	490	560	665	705	740	783	842	865	945	1380	1610	2190
Bandwidth (nm)	20	65	35	30	15	15	20	115	20	20	30	90	180
GSD (m)	60	10	10	10	20	20	20	10	20	60	60	20	20

$\implies$ B10 (recording only the cirrus information) is for atmospheric correction;

$\implies$ Different spectral bands have different spatial resolutions.

Sentinel-2 data provide multi-spectral multi-resolution images, whose analysis would not be very effective due to the lack of complete spectral patterns.

- No pixel...conventional imaging theory is no longer applicable...

Problem Statement

(a) Crop   (b) Coastal   (c) Mountainous   (d) Urban

- If we can super-resolve all the low/medium-resolution bands to achieve 10m spatial resolution, we can facilitate many practical applications:
  - (a) agricultural monitoring,
  - (b) coastal line observation,
  - (c) mountainous area modeling,
  - (d) urban area development.

- As aforementioned, data fusion, deep learning and conventional CV theory are not applicable. *Any novel SISR theory?*

- **Mathematical tools:**
  1. BCCB Matrix (Unsupervised) Fitting;
  2. Convex Self-similarity (Graph) Regularization;
  3. Large-scale Optimization.

Importance of Sentinel-2 Satellite

- Worldwide applications in military, precision agriculture, land mapping/classification, etc., across various land covers (crop, coastal, mountainous, unban areas).
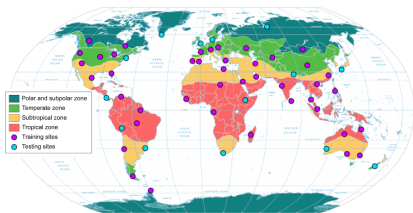
- However, lacking of pixel makes related analysis very difficult.
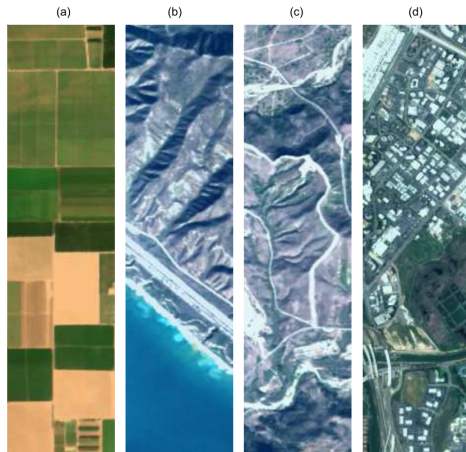
Fig. True-color composition of AVIRIS-simulated Sentinel-2 data set for (a) crop area, (b) coastal area, (c) mountainous area, and (d) urban area.

Challenges in Sentinel-2 Super-resolution

- Conventional model-based theory? *No pixels...*

- Conventional leaning-based theory? *No big data...*

- Developing novel methods is inevitable:
  - Low-rank modeling (pixel basis)
  - Unsupervised learning (and single data)
  - Other math tools:
    1. Explicit definition of self-similarity
    2. BCCB matrix theory
    3. Very large-scale optimization—ADMM

Kronecker Modeling of Sentinel-2 Data

- Sentinel-2 has $\ell = 12$ spectral bands, including $\ell_1 = 4$ 10m bands, $\ell_2 = 6$ 20m bands, and $\ell_6 = 2$ 60m bands.

- Their upsampling factors are $r_1 = 1$, $r_2 = 2$, and $r_6 = 6$, respectively.

- n_i: #(pixels) in the band with upsampling factor r_i, $\forall i \in \{1, 2, 6\}$.

- The target image $\boldsymbol{X} \in \mathbb{R}^{\ell \times n}$ (i.e., super-resolved Sentinel-2 images) contains $n \triangleq r_i^2 n_i$ high-resolution pixels.
    - Let $\boldsymbol{x}_b \in \mathbb{R}^n$ be the vectorized image of the bth band of $\boldsymbol{X}$.
    - Let $\boldsymbol{x} \triangleq [\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_\ell^T]^T \in \mathbb{R}^{\ell n}$.

- (Inconvenient) Let $\boldsymbol{Y}_b \in \mathbb{R}^{n_{i,1} \times n_{i,2}}$ ($n_{i,1} n_{i,2} = n_i$) be the bth band of $\boldsymbol{Y}$.

- (Convenient) Insert zeros to $\boldsymbol{Y}_b$, making the number of pixels also n, i.e.,

$$\boldsymbol{y}_b \triangleq \text{vec}(\boldsymbol{Y}_b \otimes (\boldsymbol{e}_1^{(r_i)} (\boldsymbol{e}_1^{(r_i)})^T) \in \mathbb{R}^n.$$

So, the data can be represented as $\boldsymbol{y} = [\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_\ell^T]^T \in \mathbb{R}^{\ell n}$.

Kronecker Modeling of Sentinel-2 Data

- The forward model:

$$\boldsymbol{y} = \boldsymbol{MBx} + \boldsymbol{n}, \qquad (2)$$

  - $\boldsymbol{M} \in \mathbb{R}^{\ell n \times \ell n}$ is block diagonal with bth diagonal block being the downsampling matrix of band b, i.e., $\boldsymbol{M}_b = \boldsymbol{M}_2^T \otimes \boldsymbol{M}_1$, where $\boldsymbol{M}_1 \triangleq \boldsymbol{I}_{n'_1/r_i} \otimes \left(\boldsymbol{e}_1^{(r_i)} (\boldsymbol{e}_1^{(r_i)})^T \right) \in \mathbb{R}^{n'_1 \times n'_1}$, and $\boldsymbol{M}_2 \triangleq \boldsymbol{I}_{n'_2/r_i} \otimes \left(\boldsymbol{e}_1^{(r_i)} (\boldsymbol{e}_1^{(r_i)})^T \right) \in \mathbb{R}^{n'_2 \times n'_2}$.

  - $\boldsymbol{B} \in \mathbb{R}^{\ell n \times \ell n}$ is also block diagonal with bth diagonal block being a block-circulant-circulant-block (BCCB) matrix, i.e., $\boldsymbol{B}_b = (\boldsymbol{F}_{n'_2} \otimes \boldsymbol{F}_{n'_1})^H \boldsymbol{\Sigma} (\boldsymbol{F}_{n'_2} \otimes \boldsymbol{F}_{n'_1})$, for spatially invariant blurring of band b (a 2D cyclic convolution associated with the point spread function (PSF) of band b).

  - $\boldsymbol{\Sigma}$ is diagonal with the diagonal vector being the vectorized version of the 2D DFT of the $n'_1 \times n'_2$ convolution/blurring kernel.

- Sentinel-2 super-resolution: to recover $\boldsymbol{x}$ from $\boldsymbol{y}$.

Problem Formulation

- To recover $\boldsymbol{x}$ from $\boldsymbol{y}$, the data-fitting is naturally $\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{MBx}\|^2$.

- To handle the ill-posed inverse problem, we propose two strategies:

  1. **Low-rank modeling**
     - More than 99% of the signal energy of the 12-bands Sentinel-2 image is retained in the $p = 5$ principal components.
     - Therefore, we can write $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell]^T = \boldsymbol{UZ}$, or equivalently

     $$\boldsymbol{x} = (\boldsymbol{U} \otimes \boldsymbol{I}_n)\text{vec}(\boldsymbol{Z}^T), \qquad (3)$$

     for some full column rank $\boldsymbol{U} \in \mathbb{R}^{\ell \times p}$. $\boldsymbol{Z}$ is termed eigenimage.

  2. **Regularization**
     - From (2) and (3), a regularized criterion is proposed:

     $$\min_{\boldsymbol{z}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{MB}(\boldsymbol{U} \otimes \boldsymbol{I}_n)\boldsymbol{z}\|_2^2 + \lambda\phi(\boldsymbol{z}), \qquad (4)$$

     where $\boldsymbol{z} \triangleq \text{vec}(\boldsymbol{Z}^T)$, $\lambda > 0$, and ϕ is a regularizer to further mitigate the ill-posedness.

Self-similarity Regularization

- *As some spatial information is missing (e.g., blurring), just find other spatial information as a compensation!*

- Self-similarity: An important spatial structural information commonly observed in natural images, including RGB, SAR, MRI, multispectral, hyperspectral images, etc.

- No explicit definition...

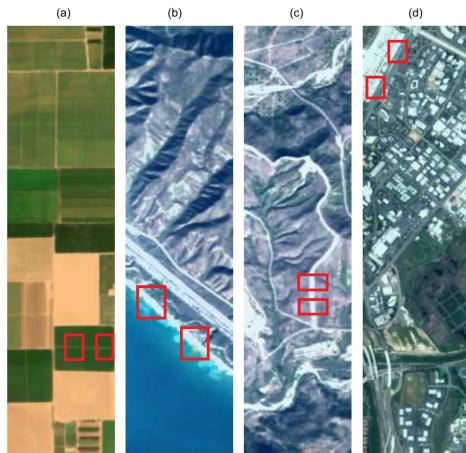- Plug-and-play in machine learning has no convergence guarantee...

Fig. True-color composition of AVIRIS-simulated Sentinel-2 data set for (a) crop area, (b) coastal area, (c) mountainous area, and (d) urban area.

Unsupervised Learning: **Explicit, Convex** and **Scene-adapted** Self-similarity Regularizer

- **Self-similarity:** existence of similar patches at different locations in a given image *(critical prior information for the ill-posed inverse problem)*.

- A widely observed phenomenon in natural images, including optical and radar images [Zhao'14], [Deledalle'14], but there is no explicit mathematical definition.

- For the first time, we define it as a convex regularization function of the eigenimage $\boldsymbol{z}$:

$$\phi(\boldsymbol{z}) \triangleq \frac{1}{2} \sum_{b=1}^{p} \sum_{(i,j) \in \mathcal{K}} \alpha_{i,j} \left\| \boldsymbol{P}_i \boldsymbol{z}_b - \boldsymbol{P}_j \boldsymbol{z}_b \right\|_2^2. \tag{5}$$

Why convexity? (fast software with convergence guarantee)

- The self-similarity graph $\mathcal{K}$ is scene-adapted (i.e., automatically learned from the Sentinel-2 image itself), rather than fixed (cf. TV regularization).

- How to unsupervisedly learn the self-similarity pattern $(\mathcal{K}, \alpha_{i,j})$?

[Zhao'14] Y. Zhao et al., "Hyperspectral imagery superresolution by spatial-spectral joint nonlocal similarity," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014.
[Deledalle'14] C.-A. Deledalle et al., "Exploiting patch similarity for SAR image processing: the nonlocal paradigm," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 69-78, 2014.

Regularizer using Self-similarity Graph
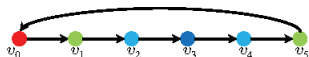
(a)

(b)
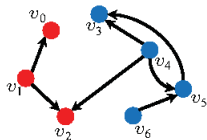
(c)
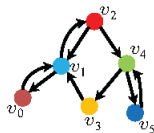
(d)

- GSP: a newly emerged area in signal processing community for natural language understanding.

- We explicitly define self-similarity as a "weighted" graph $\mathcal{K}$:
  - vertex: patch;
  - edge: connecting similar patches;
  - weight: similarity measure $\alpha_{i,j} \geq 0$.

The graph is transformed into the aforementioned convex algebra (5), allowing the adoption of distributed optimization theory with convergence guarantee.

Convex Criterion for Sentinel-2 SISR

- A convex criterion has been designed as follows:

$$\min_{\boldsymbol{z}, \boldsymbol{v}, \boldsymbol{V}_{i,j}} \frac{\|\boldsymbol{y} - \boldsymbol{M}\boldsymbol{B}\boldsymbol{v}\|_2^2}{\lambda} + \sum_{(i,j)\in\mathcal{K}} \alpha_{i,j} \|(\boldsymbol{P}_i - \boldsymbol{P}_j)\boldsymbol{V}_{i,j}\|_F^2$$

$$\text{s.t.} \quad \boldsymbol{v} = (\boldsymbol{U} \otimes \boldsymbol{I}_n)\boldsymbol{z},$$

$$\boldsymbol{V}_{i,j} = \boldsymbol{Z}^T, \ \forall (i,j) \in \mathcal{K},$$

- The acquired Sentinel-2 image $\boldsymbol{y}$ is modeled as the blurred ($\boldsymbol{B}$) and downsampled ($\boldsymbol{M}$) version of the super-resolved image $\boldsymbol{v}$ — data fidelity!

- $\boldsymbol{U}$: feature subspace (also automatically learned).

A large-scale optimization with BCCB structure.

Convex Optimization for Signal Processing and Communications

From Fundamentals to Applications

Chong-Yung Chi · Wei-Chiang Li · Chia-Hsiang Lin

CRC Press
Taylor & Francis Group

BCCB structure of blurring matrix B

- The BCCB structure is employed to design fast algorithm behind our software: *Scene-adapted Self-similarity regularized Sentinel-2 Super-resolution* algorithm!

Algorithm 1

1: **Given** (M, B, U, y), $(\mathcal{K}, \alpha_{i,j})$, $\lambda > 0$ and $\mu > 0$.
2: Initialize $z^0 := 0_{pn}$, $d^0 := 0_{\ell n}$, $D_{i,j}^0 := 0_{n \times p}$ (or warm start). Set $k = 0$.
3: **repeat**
4:     Update $(v^{k+1}, \{V_{i,j}^{k+1}\})$ $\in$
    $\arg\min_{v, \{V_{i,j}\}} \mathcal{L}(v, \{V_{i,j}\}, z^k, d^k, \{D_{i,j}^k\})$;
5:     Update
    $z^{k+1} \in \arg\min_z \mathcal{L}(v^{k+1}, \{V_{i,j}^{k+1}\}, z, d^k, \{D_{i,j}^k\})$;
6:     Update $d^{k+1} := d^k + v^{k+1} - (U \otimes I_n)z^{k+1}$ and
    $D_{i,j}^{k+1} := D_{i,j}^k + V_{i,j}^{k+1} - (Z^{k+1})^T$;
7:     $k := k + 1$;
8: **until** the predefined stopping criterion is met.
9: **Output** the super-resolved Sentinel-2 image $\hat{x} \equiv$
    $\mathrm{vec}(\hat{X}^T) := (U \otimes I_n)z^k$ (cf. (2)).

Theorem 1 *Consider any graph $\mathcal{K} \subseteq \{(i,j) \mid i, j \in \mathcal{I}_N, i \neq j\}$ and $\alpha_{i,j} \geq 0$, $\forall (i,j) \in \mathcal{K}$, where N is the number of patches. The sequence $\{z^k\}$ generated by Algorithm 1 converges to an optimal solution of the self-similarity regularized problem.*

1. Closed-form solutions are derived for all the algorithmic steps $\implies$ fast!
2. Theoretical guarantee of convergence to global optimum of our learning criterion!

- (Line 4 in Algorithm)

$$v^{k+1} \in \arg\min_v \frac{1}{2}\|y - MBv\|_2^2 + \frac{\mu}{2}\|(U \otimes I_n)z^k - v - d^k\|_2^2, \quad (123)$$

$$V_{i,j}^{k+1} \in \arg\min_{V_{i,j}} \frac{\lambda\alpha_{i,j}}{2}\|(P_i - P_j)V_{i,j}\|_F^2 + \frac{\mu}{2}\|(Z^k)^T - V_{i,j} - D_{i,j}^k\|_F^2. \quad (124)$$

1. (Solution for (123)) $\quad v_b^{k+1} = (B_b^T M_b^T M_b B_b + \mu I_n)^{-1}(B_b^T M_b^T y_b + \mu \bar{v}_b).$ inversion of a huge $n \times n$ matrix

- **Proposition 5** Let $M' \triangleq I_{n'/r} \otimes \left(e_1^{(r)}(e_1^{(r)})^T\right) \in \mathbb{R}^{n' \times n'}$, and $F_{n'} \in \mathbb{C}^{n' \times n'}$ be the DFT matrix. Then,

$$F_{n'}M'F_{n'}^H = F_{n'}^H M'F_{n'} = \frac{1}{r}(1_r \otimes I_{n'/r})(1_r^T \otimes I_{n'/r}).$$

- **Proposition 6** The following identity holds true for matrices with proper dimensionality:

$$(A + UBV)^{-1} = A^{-1} - A^{-1}U(B^{-1} + VA^{-1}U)^{-1}VA^{-1}. \quad \text{Woodbury matrix inversion lemma}$$

- $B_b^T M_b^T M_b B_b = F^H \Sigma^H F(M_2 \otimes M_1)F^H \Sigma F = F^H \Sigma^H (F_{n_2'}M_2 F_{n_2'}^H \otimes F_{n_1'}M_1 F_{n_1'}^H)\Sigma F$

$\qquad\qquad\qquad = F^H \Sigma^H D \frac{1}{r_?^2}D^H \Sigma F,$

- $(B_b^T M_b^T M_b B_b + \mu I_n)^{-1} = \frac{1}{\mu}I_n - \frac{1}{\mu^2}F^H \Sigma^H D\left(r_i^2 I_{n_i} + \frac{1}{\mu}D^H|\Sigma^2|D\right)^{-1}D^H \Sigma F.$

- Remarkably, $\left(r_i^2 I_{n_i} + \frac{1}{\mu}D^H|\Sigma^2|D\right)$ is diagonal (making its inversion easily computable).

Fact 2 (a) If each row of A_k ($k \in \{1, 2\}$) has maximally one nonzero entry, then each row of $A_1 \otimes A_2$ also has maximally one nonzero entry.

(b) If A_1 is diagonal, and each row of A_2 has maximally one nonzero entry, then $A_2^H A_1 A_2$ must be diagonal.

S4: High-dimensional Matrix Inversion II — Column Sparsity Pattern

- (Line 4 in Algorithm)

$$v^{k+1} \in \arg\min_{v} \frac{1}{2}\|y - MBv\|_2^2 + \frac{\mu}{2}\|(U \otimes I_n)z^k - v - d^k\|_2^2, \quad (123)$$

$$V_{i,j}^{k+1} \in \arg\min_{V_{i,j}} \frac{\lambda\alpha_{i,j}}{2}\|(P_i - P_j)V_{i,j}\|_F^2 + \frac{\mu}{2}\|(Z^k)^T - V_{i,j} - D_{i,j}^k\|_F^2. \quad (124)$$

2. (Solution for (124)) Let $v_{i,j}^{(c)}$ be the cth column of $V_{i,j}^{k+1}$. Then, (124) can be decoupled:

$$\min_{v_{i,j}^{(c)}} \frac{\lambda\alpha_{i,j}}{2}\|(P_i - P_j)v_{i,j}^{(c)}\|_2^2 + \frac{\mu}{2}\|v_{i,j}^{(c)} - \delta_{i,j}^{(c)}\|_2^2, \quad (130)$$

where $\delta_{i,j}^{(c)}$ denotes the cth column of $(Z^k)^T - D_{i,j}^k$. $\boxed{\text{inversion of a huge } n \times n \text{ matrix}}$

- Note that since $(P_i - P_j) \in \mathbb{R}^{q^2 \times n}$ (the patch size is typically set as $q \times q \equiv 6 \times 6$).
- $(P_i - P_j) \in \mathbb{R}^{q^2 \times n}$ is actually a column-sparse matrix. Then, we can reformulate (130) as

$$\min_{[v_{i,j}^{(c)}]_{\mathcal{I}},[v_{i,j}^{(c)}]_{\bar{\mathcal{I}}}} \frac{\lambda\alpha_{i,j}}{2}\|[P_i - P_j]_{\mathcal{I}}[v_{i,j}^{(c)}]_{\mathcal{I}}\|_2^2 + \frac{\mu}{2}\|[v_{i,j}^{(c)}]_{\mathcal{I}} - [\delta_{i,j}^{(c)}]_{\mathcal{I}}\|_2^2 + \frac{\mu}{2}\|[v_{i,j}^{(c)}]_{\bar{\mathcal{I}}} - [\delta_{i,j}^{(c)}]_{\bar{\mathcal{I}}}\|_2^2, \quad (131)$$

- A computationally efficient closed-form solution is hence obtained as follows:

$$[v_{i,j}^{(c)}]_{\mathcal{I}}^{\star} = \left[[P_i - P_j]_{\mathcal{I}}^T[P_i - P_j]_{\mathcal{I}} + \frac{\mu I_{|\mathcal{I}|}}{\lambda\alpha_{i,j}}\right]^{-1}\left(\frac{\mu[\delta_{i,j}^{(c)}]_{\mathcal{I}}}{\lambda\alpha_{i,j}}\right),$$

$$[v_{i,j}^{(c)}]_{\bar{\mathcal{I}}}^{\star} = [\delta_{i,j}^{(c)}]_{\bar{\mathcal{I}}}.$$

- Note that we now only need to compute the $\boxed{\text{inverse of a small } |\mathcal{I}| \times |\mathcal{I}| \text{ matrix}}$ with $|\mathcal{I}| \leq 2q^2$.

- (Line 5 in Algorithm 11)

  The problem can be first simplified as

  $$z^{k+1} \in \arg\min_{z} \|(U \otimes I_n)z - v^{k+1} - d^k\|_2^2 + \sum_{(i,j)\in\mathcal{K}} \|Z^T - V_{i,j}^{k+1} - D_{i,j}^k\|_F^2. \tag{132}$$

  – Both terms are quadratic terms of z. By defining $\Psi_0 \triangleq (U \otimes I_n)^T (U \otimes I_n) + |\mathcal{K}|I_{pn}$ and $\Psi_2 \triangleq \sum_{(i,j)\in\mathcal{K}} \left(V_{i,j}^{k+1} + D_{i,j}^k\right)$, the closed-form solution of (132) can be derived as

  $$z^{k+1} = \Psi_0^{-1} \left[(U \otimes I_n)^T (v^{k+1} + d^k) + \mathrm{vec}(\Psi_2)\right]. \tag{133}$$

  – We do not need to compute $\Psi_0^{-1} \in \mathbb{R}^{pn \times pn}$. $\boxed{\text{inversion of a huge } pn \times pn \text{ matrix}}$

  – By further defining $\Psi_1 \triangleq \mathrm{vec}_{n\times\ell}^{-1}(v^{k+1} + d^k)\, U$, the closed-form solution (133) can be refined as

  $$z^{k+1} = \Psi_0^{-1} \left[(U \otimes I_n)^T (v^{k+1} + d^k) + \mathrm{vec}(\Psi_2)\right]$$
  $$= \mathrm{vec}\left((\Psi_1 + \Psi_2)\left(U^T U + |\mathcal{K}|I_p\right)^{-1}\right). \tag{134}$$

  – Note that (134) is much more computationally efficient than (133) as the involved matrix inversion associates with a $\boxed{\text{much smaller } p \times p \text{ matrix}}$.

Experimental Setting

- Datasets:
    1. real Sentinel-2 dataset: 4 most commonly studied scenes (corresponding to the 4 aforementioned applications: agricultural monitoring, coastal line observation, mountainous area modeling, urban area development),
    2. synthetic dataset: for performance comparison only.

- Key peer methods:
    1. area-to-point regression kriging (ATPRK) [Wang'16],
    2. multi-resolution sharpening approach (MUSA) [Paris'18].

- Performance measures:
    1. spectral angle mapper (SAM) $\in [-\pi, \pi]$ (the smaller, the better),
    2. root-mean-square error (RMSE) $\in [0, \infty)$ (the smaller, the better),
    3. signal-to-reconstruction error (SRE) $\in (-\infty, \infty)$ (the larger, the better),
    4. universal image quality index (UIQI) $\in [-1, 1]$ (the larger, the better),

    measuring spectral distortion, loss of correlation, luminance distortion and contrast distortion.

[Wang'16] Q. Wang et al., "Fusion of Sentinel-2 images," *Remote Sensing of Environment*, 2016.
[Paris'18] C. Paris et al., "A novel sharpening approach for super-resolving multi-resolution optical images," *IEEE Trans. Geoscience and Remote Sensing*, 2018.
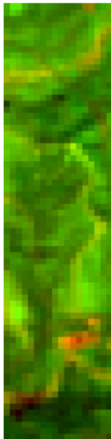
Real Sentinel-2 Data 1: Mountainous Area

- (True color composition — 10m)   B4(red) + B3(green) + B2(blue)   [ref. image]
- (False color composition — 20m)   B5+B6+B7   &   B8a+B11+B12
- (False color composition — 60m)   B1+B9

(a) RGB (10m) (b) B1–9 (60m) (c) B1–9 (SR) (d) B5–6–7 (20m) (e) B5–6–7 (SR) (f) B8a–11–12 (20m) (g) B8a–11–12 (SR)
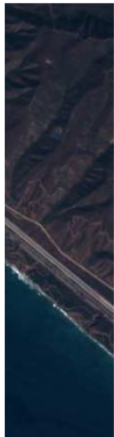
- (True color composition — 10m)   B4(red) + B3(green) + B2(blue)   [ref. image]
- (False color composition — 20m)   B5+B6+B7   &   B8a+B11+B12
- (False color composition — 60m)   B1+B9

(a) RGB (10m)   (b) B1–9 (60m)   (c) B1–9 (SR)   (d) B5–6–7 (20m)   (e) B5–6–7 (SR)   (f) B8a–11–12 (20m)   (g) B8a–11–12 (SR)
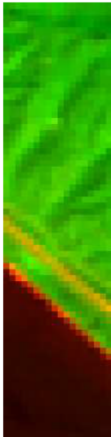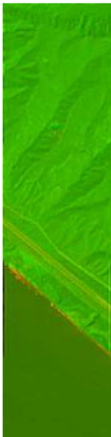
Real Sentinel-2 Data 3: Crop Area

- (True color composition — 10m)   B4(red) + B3(green) + B2(blue)   [ref. image]
- (False color composition — 20m)   B5+B6+B7   &   B8a+B11+B12
- (False color composition — 60m)   B1+B9

(a) RGB (10m)   (b) B1–9 (60m)   (c) B1–9 (SR)   (d) B5–6–7 (20m)   (e) B5–6–7 (SR)   (f) B8a–11–12 (20m)   (g) B8a–11–12 (SR)

- (True color composition — 10m)   B4(red) + B3(green) + B2(blue)   [ref. image]
- (False color composition — 20m)   B5+B6+B7   &   B8a+B11+B12
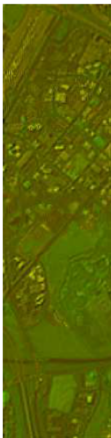- (False color composition — 60m)   B1+B9
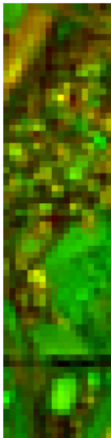
(a) RGB (10m) (b) B1–9 (60m) (c) B1–9 (SR) (d) B5–6–7 (20m) (e) B5–6–7 (SR) (f) B8a–11–12 (20m) (g) B8a–11–12 (SR)

Quntitative Comparison on Synthetic Data

Our S4 software:

- best spectral property (joint learning) — key to multi-resolution analysis,
- best global accuracy (or close to the best),
- much faster than key competitors.

Datasets	Methods	SAM (degree)	RMSE	SRE (dB)	UIQI	T (sec.)
Crop Dataset	MUSA	0.9518	55.5026	32.6243	0.9853	501.13
	ATPRK	1.2341	71.8985	31.5727	0.9819	226.32
	DSen2	2.4042	139.5301	28.2986	0.9442	**10.08**
	S2Sharp	1.9193	152.2005	26.9332	0.9530	20.78
	SSSS	**0.8040**	**50.0046**	**33.1973**	**0.9864**	100.05
Coastal Dataset	MUSA	1.8156	24.7417	27.9673	**0.9377**	455.17
	ATPRK	1.8678	21.1794	28.9295	0.9214	222.48
	DSen2	4.5553	57.5471	18.5766	0.9161	**10.23**
	S2Sharp	2.3934	47.7705	18.3089	0.9191	20.42
	SSSS	**1.1622**	**18.0857**	**29.3141**	0.9317	105.83
Mountainous Dataset	MUSA	1.0779	17.6223	31.3936	0.9915	400.62
	ATPRK	1.7754	25.6363	29.0110	0.9878	226.90
	DSen2	2.7074	48.3057	19.1132	0.9577	**10.20**
	S2Sharp	2.4254	51.5264	19.0960	0.9401	22.46
	SSSS	**1.0478**	**17.1165**	**31.6545**	**0.9916**	97.99
Urban Dataset	MUSA	1.5680	54.8530	27.2624	0.9727	425.72
	ATPRK	1.6179	47.1817	28.5205	**0.9885**	222.04
	DSen2	2.8366	108.9142	22.3203	0.9371	**9.77**
	S2Sharp	2.5630	102.3011	22.3814	0.9528	24.13
	SSSS	**1.0553**	**35.4103**	**30.0981**	0.9862	101.72

Summary

- ESA's Sentinel-2 satellite is important for various Earth observation missions, but effective analysis is hampered by its multi-resolution nature (10/20/60 meters).

- Super-resolving the 20/60m bands to achieve 10m resolution is desired, and our S4 software achieves so by incorporating the self-similarity prior info. (to compensate the loss of spatial info.) into a convex optimization framework.

- S4 is capable of reconstructing spatial details in various scenes, while preserving spectral characteristics, with source codes released on *IEEE Code Ocean* and my website (cf. **QR Code**).

Thank You for Your Attention.

Special thanks to my first batch of students!

My First Batch of Students:
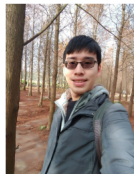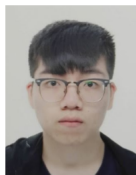

Tzu-Hsuan Lin

Cheng-Yu Sie

Yen-Cheng Lin

Chi-Hung Kao

Po-Wei Tang

Yangrui Liu

Jhao-Ting Lin

Man-Chun Chu

Zi-Chao Leng

Yi-Hsun Lee

Pang-Yu Lin

*Recruiting
(Postdoc, RA, etc.)*

蒐集大數據無需、
訓練模型也不需。
若要得出其細節、
予一組數據足歟！

平平仄仄仄平平
仄仄平平仄仄平
仄仄平平平仄仄
平平仄仄仄平平

Thank You for Your Attention.

Q & A

"No Magic, Only Basic."