HI, BCD! HYBRID INEXACT BLOCK COORDINATE DESCENT FOR HYPERSPECTRAL SUPER-RESOLUTION

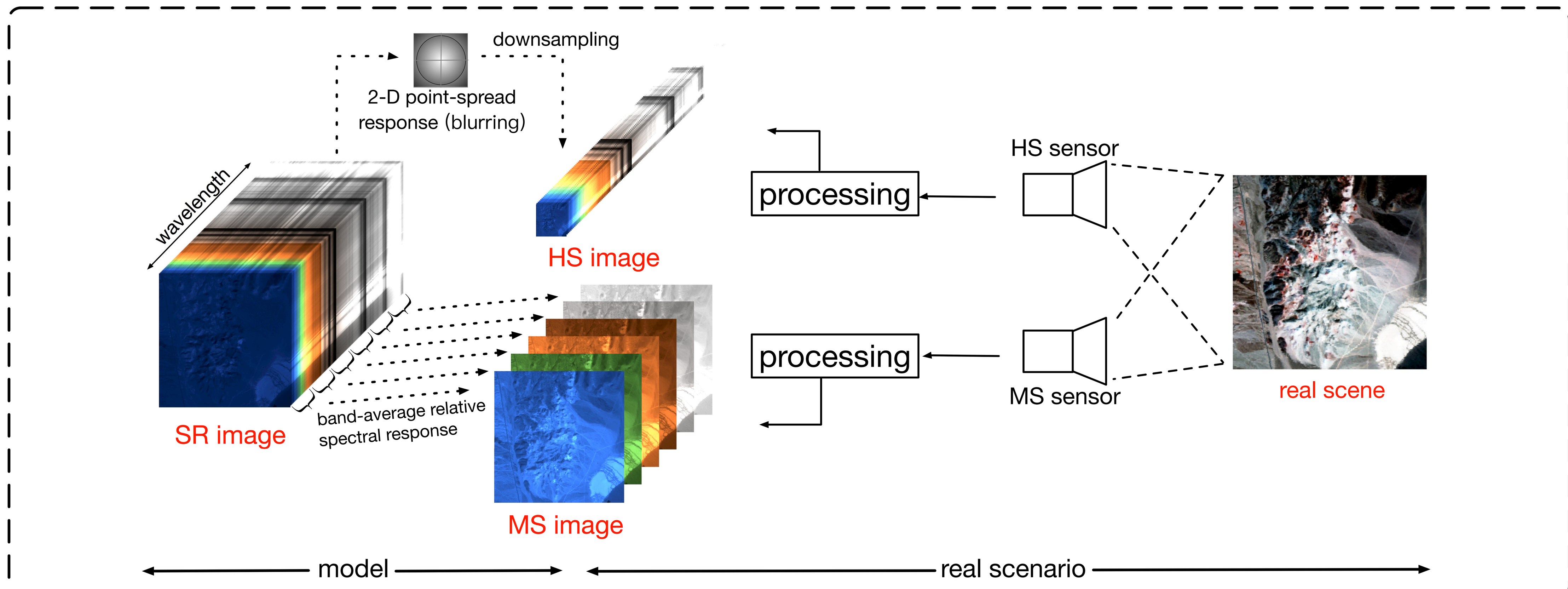
†Ruiyuan Wu, †Chun-Hei Chan, ‡Hoi-To Wai, †Wing-Kin Ma, and *Xiao Fu

†Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China
‡School of ECEE, Arizona State Univ., AZ, USA, * School of EECS, Oregon State Univ., OR, USA

†* Supported in part by the Hong Kong RGC GRF under Project ID 14205414, and in part by the U.S. National Science Foundation under project NSF ECCS-1608961

Key Points

- We tackle the hyperspectral super-resolution problem using matrix factorization and first-order optimization.
- We devise a novel inexact block coordinate descent method which employs hybrid proximal gradient and Frank-Wolfe updates.



Hyperspectral Super-Resolution (HSR)

- **Spectral sensors:** capture scenes in multiple spectral bands
  - Hyperspectral (HS) sensors
    - specifications:

Sensor	AVIRIS	HYPERION	HJ-1A
Band number	224	242	128
Wavelength range	0.4 - 2.5μm	0.4 - 2.5μm	0.45-0.95μm
Spatial resolution	20 m	30 m	80 m

    - an HS image has low-spatial and high-spectral resolution;
  - Multispectral (MS) sensors
    - specifications:

Sensor	QuickBird	WorldView-2
Band number	4	8
Wavelength range	0.4 - 0.9μm	0.4 - 1μm
Spatial resolution	2.16 m	1.85 m

    - an MS image has high-spatial and low-spectral resolution.
  - Super-resolution (SR), high spatial-spectral resolution, sensors? Not exist.
- **HSR:** recover an SR image from an HS-MS image pair.
- **Applications:** high-spatial-resolution mapping of, e.g., minerals, urban surface materials, plant species, etc.

Problem Statement

- **Signal model:**

$$\text{MS image model: } \boldsymbol{Y}_M = \boldsymbol{F}\boldsymbol{X} + \boldsymbol{V}_M$$
$$\text{HS image model: } \boldsymbol{Y}_H = \boldsymbol{X}\boldsymbol{G} + \boldsymbol{V}_H$$

  - $\boldsymbol{X} \in \mathbb{R}^{M \times L}$ is spectral-spatial matrix of the SR image;
  - $\boldsymbol{Y}_H \in \mathbb{R}^{M \times L_H}$ is the spectral-spatial matrix of the HS image;
  - $\boldsymbol{Y}_M \in \mathbb{R}^{M_M \times L}$ is the spectral-spatial matrix of the MS image;
  - $\boldsymbol{F} \in \mathbb{R}^{M_M \times M}$ is the spectral degradation matrix;
  - $\boldsymbol{G} \in \mathbb{R}^{L \times L_H}$ is the spatial degradation matrix;
  - $\boldsymbol{V}_M$ and $\boldsymbol{V}_M$ are noise.

- **Assumption:** the SR image has low rank, i.e., $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S}$
  - $N \ll \min\{M, L\}$;
  - $\boldsymbol{A}$ and $\boldsymbol{S}$ follow the linear mixture model
    - it is widely-used in remote sensing;
    - $\boldsymbol{A} \in \mathcal{A}$, where $\mathcal{A} = [0,1]^{M \times N}$;
    - $\boldsymbol{S} \in \mathcal{S}$, where $\mathcal{S} = \{\boldsymbol{S} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L] | \boldsymbol{s}_i \geq 0, \boldsymbol{1}^T\boldsymbol{s}_i = 1, \forall\, i\}$.
- **Structured matrix factorization (SMF) formulation:**

$$\min_{\boldsymbol{A}\in\mathcal{A}, \boldsymbol{S}\in\mathcal{S}} f(\boldsymbol{A}, \boldsymbol{S}) := \frac{1}{2}\|\boldsymbol{Y}_M - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_F^2 + \frac{1}{2}\|\boldsymbol{Y}_H - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_F^2$$

  - **Challenge:** the number of unknowns is very large
    - L, the pixel number, may range from thousands to millions;
  - **Aim:** develop a computational efficient optimization scheme.

Exact Block Coordinate Descent (BCD)

- Exact BCD works by recursively solving

$$\boldsymbol{S}^{k+1} = \arg\min_{\boldsymbol{S}\in\mathcal{S}} f(\boldsymbol{A}^k, \boldsymbol{S})$$
$$\boldsymbol{A}^{k+1} = \arg\min_{\boldsymbol{A}\in\mathcal{A}} f(\boldsymbol{A}, \boldsymbol{S}^{k+1})$$

  - It guarantees convergence to a stationary point.
- FUMI [1] is the state-of-the-art exact BCD-based algorithm
  - It uses customize-designed, ADMM-based, solvers for $\boldsymbol{A}$ and $\boldsymbol{S}$.
- **Limitation:** computationally expensive when data sizes are large.

Proposed Inexact BCD Scheme

- **Scheme I**: inexact BCD by proximal gradient (PG)

$$\boldsymbol{S}^{k+1} = \Pi_{\mathcal{S}}(\boldsymbol{S}^k - \gamma_{S,k}\nabla_{\boldsymbol{S}} f(\boldsymbol{A}^k, \boldsymbol{S}^k))$$
$$\boldsymbol{A}^{k+1} = \Pi_{\mathcal{A}}(\boldsymbol{A}^k - \gamma_{A,k}\nabla_{\boldsymbol{A}} f(\boldsymbol{A}^k, \boldsymbol{S}^{k+1}))$$

  - $\Pi_{\mathcal{S}}(\cdot)$ and $\Pi_{\mathcal{A}}(\cdot)$ are the projections onto $\mathcal{S}$ and $\boldsymbol{A}$, resp.
    - $\Pi_{\mathcal{S}}(\cdot)$ is a column-wise unit simplex projection;
    - $\Pi_{\mathcal{A}}(\cdot) = \max\{0, \min\{1, \cdot\}\}$ (elementwise operations);
  - $\gamma_{S,k}$ and $\gamma_{A,k}$ are the step sizes;
  - An instance of BSUM [2];
  - Per-iteration complexities:

$\boldsymbol{S}$-update	$\mathcal{O}(LN(M + \log(N)) + N \cdot \mathrm{nnz}(\boldsymbol{G}) + N^2M)$
$\boldsymbol{A}$-update	$\mathcal{O}(LNM + N \cdot \mathrm{nnz}(\boldsymbol{G}) + N^2L)$

- **Scheme II:** inexact BCD by Frank-Wolfe (FW)

$$\boldsymbol{S}^{k+1} = \boldsymbol{S}^k + \alpha_{S,k}(\boldsymbol{P}_S^k - \boldsymbol{S}^k)$$
$$\boldsymbol{A}^{k+1} = \boldsymbol{A}^k + \alpha_{A,k}(\boldsymbol{P}_A^k - \boldsymbol{A}^k)$$

  - It is projection free;
  - $\boldsymbol{P}_S^k$ and $\boldsymbol{P}_A^k$ are the FW directions of $\boldsymbol{S}$ and $\boldsymbol{A}$, resp.
    - $\boldsymbol{P}_S^k = \arg\min_{\boldsymbol{Z}\in\mathcal{S}} \langle \nabla_{\boldsymbol{S}} f(\boldsymbol{A}^k, \boldsymbol{S}^k), \boldsymbol{Z}\rangle$
    - $\boldsymbol{P}_A^k = \arg\min_{\boldsymbol{Z}\in\mathcal{A}} \langle \nabla_{\boldsymbol{A}} f(\boldsymbol{A}^k, \boldsymbol{S}^{k+1}), \boldsymbol{Z}\rangle$
  - $\alpha_{S,k}$ and $\alpha_{A,k}$ are the step sizes;
  - An instance of CBCG [3] (roughly speaking).
  - Per-iteration complexities:

$\boldsymbol{S}$-update	$\mathcal{O}(LNM + N \cdot \mathrm{nnz}(\boldsymbol{G}))$
$\boldsymbol{A}$-update	$\mathcal{O}(LNM + N \cdot \mathrm{nnz}(\boldsymbol{G}))$

- **Scheme III:** hybrid inexact BCD (HiBCD)

$$\boldsymbol{S}^{k+1} = \mathtt{UD}_S(\boldsymbol{A}^k, \boldsymbol{S}^k)$$
$$\boldsymbol{A}^{k+1} = \mathtt{UD}_A(\boldsymbol{A}^k, \boldsymbol{S}^{k+1})$$

  - $\mathtt{UD}_S$ and $\mathtt{UD}_A$ can be either the PG or FW updates, e.g.,
    - PG update for $\boldsymbol{A}$, FW update for $\boldsymbol{S}$;
  - **convergence result:**

Theorem 1

The HiBCD scheme guarantees convergence to a stationary point of the SMF for HSR. Also, its convergence rate, measured by means of the FW gap, is $\mathcal{O}(1/\sqrt{k})$.

  - this result applies to a wide class of optimization problems.

Simulations

- **Algorithms under comparison**
  - PGiBCD: PG update for both $\boldsymbol{A}$ and $\boldsymbol{S}$;
  - FWiBCD: FW update for both $\boldsymbol{A}$ and $\boldsymbol{S}$;
  - HiBCD: PG update for $\boldsymbol{A}$, FW update for $\boldsymbol{S}$;
  - FUMI: the state-of-the-art exact BCD algorithm.
- **Synthetic data experiment**
  - Settings:
    - $N = 9$, $L = 100^2$, $M = 224$, $L_H = 25^2$, $M_M = 6$;
    - $\boldsymbol{A}$ is from the USGS digital spectral library;
    - $\boldsymbol{S}$ is from an abundance map of the AVIRIS Cuprite dataset;
    - $\boldsymbol{G}$ corresponds to 11×11 Gaussian point spreading with variance $\sigma^2 = 1.7^2$, followed by downsampling with ratio 4;
    - $\boldsymbol{F}$ corresponds to the LANDSAT specification;
  - Averaged performance over 100 independent trials

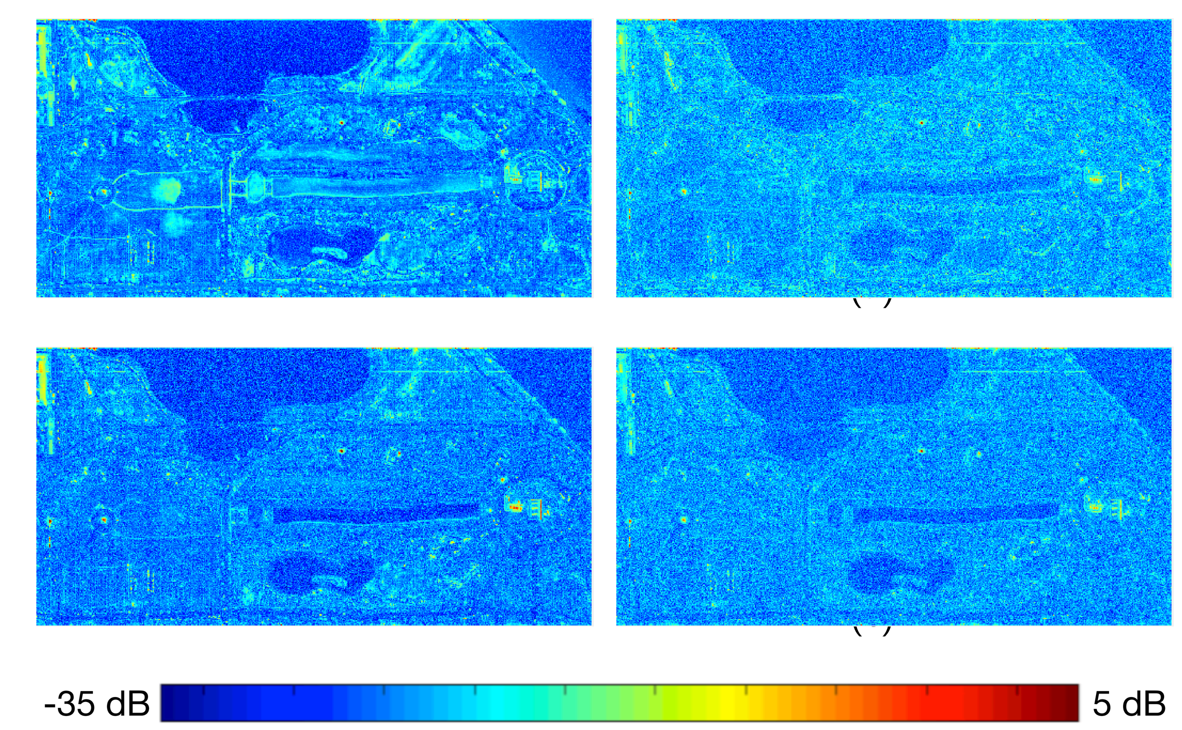
SNR	Method	Runtime (sec.)	Iterations	PSNR (dB)
20	FUMI	8.31 ± 1.75	214.44 ± 51.59	16.39 ± 0.41
	PGiBCD	3.41 ± 0.43	503.04 ± 62.81	$\mathbf{17.68 \pm 0.51}$
	FWiBCD	$\mathbf{1.01 \pm 0.09}$	$\mathbf{164.30 \pm 13.55}$	17.59 ± 0.49
	HiBCD	1.59 ± 0.18	272.40 ± 32.08	17.66 ± 0.51
30	FUMI	15.80 ± 4.48	435.70 ± 125.92	22.72 ± 0.81
	PGiBCD	7.01 ± 0.96	1057.14 ± 140.57	$\mathbf{24.33 \pm 0.89}$
	FWiBCD	$\mathbf{1.44 \pm 0.20}$	$\mathbf{237.12 \pm 34.17}$	24.19 ± 0.85
	HiBCD	2.51 ± 043	434.34 ± 75.09	24.27 ± 0.87
40	FUMI	19.45 ± 5.89	566.60 ± 181.69	32.41 ± 0.98
	PGiBCD	14.78 ± 3.25	2235.18 ± 496.21	$\mathbf{33.02 \pm 1.08}$
	FWiBCD	$\mathbf{3.11 \pm 0.53}$	$\mathbf{515.12 \pm 87.38}$	32.61 ± 1.04
	HiBCD	5.29 ± 1.07	932.14 ± 191.53	32.72 ± 1.05

- **Semi-real data experiment**
  - Settings:
    - $N = 30$, $L = 520 \times 260$, $M = 191$, $L_H = 130 \times 65$, $M_M = 6$;
    - $\boldsymbol{X}$ is cropped from the HYDICE Washington DC dataset;

  - Averaged performance over 50 trials (SNR: 40dB)

Method	Runtime (sec.)	Iterations	PSNR (dB)
FUMI	1162.53 ± 235.89	$\mathbf{950.23 \pm 194.31}$	41.24 ± 0.53
PGiBCD	560.64 ± 18.84	2115.23 ± 71.34	$\mathbf{46.88 \pm 0.04}$
FWiBCD	$\mathbf{304.73 \pm 9.71}$	1610.47 ± 51.01	41.34 ± 0.12
HiBCD	310.52 ± 8.35	1689.94 ± 46.20	44.25 ± 0.09

  - an instance of the mean square error maps of the algorithms



MSE maps of (a) PGiBCD, (b) FWiBCD, (c) HiBCD and (d) FUMI

Conclusion

- A hybrid inexact BCD scheme was proposed for HSR.
- Computational and convergence issues were dealt with.
- Numerical results showed promising runtime performance.

Reference

[1] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill, "Multiband image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, 2016.

[2] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[3] A. Beck, E. Pauwels, and S. Sabach, "The cyclic block conditional gradient method for convex optimization problems," *SIAM J. Optim.*, vol. 25, no. 4, pp. 2024–2049, 2015.

香港中文大學
The Chinese University of Hong Kong