M.Phil Oral Defense

Hyperspectral Super-resolution Using Matrix Factorization

Chan Chun Hei

Supervisor: Prof. Wing-Kin Ma

5th Sept. 2017

DSP-ST Laboratory
Department of Electronic Engineering
The Chinese University of Hong Kong

Outline

Introduction and Motivation
- Spectral Signature and Spectral Image
- Applications
- The Instruments
- What is the Problem ?

Hyperspectral Super-resolution (HSR)
- The HSR Problem
- State of the Arts

Fast Algorithm Framework of HSR (proposed)
- Preliminaries
- A Very Fast Algorithm Framework

Experiments

Summary

Appendices and References

Spectral Signature

- Sun light possesses large amount of energy in the EM spectrum.
- Material partially *absorbs* the energy from the spectrum and *reflects* the remaining energy $\rightarrow$ spectral signature.

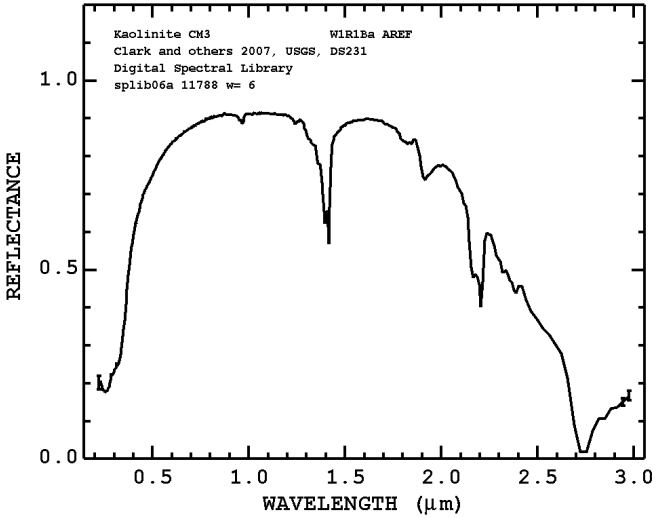


Figure: Spectral signature of Kaolinite within $0.4\,\mu$m and $3.0\,\mu$m [1].

Spectral Image

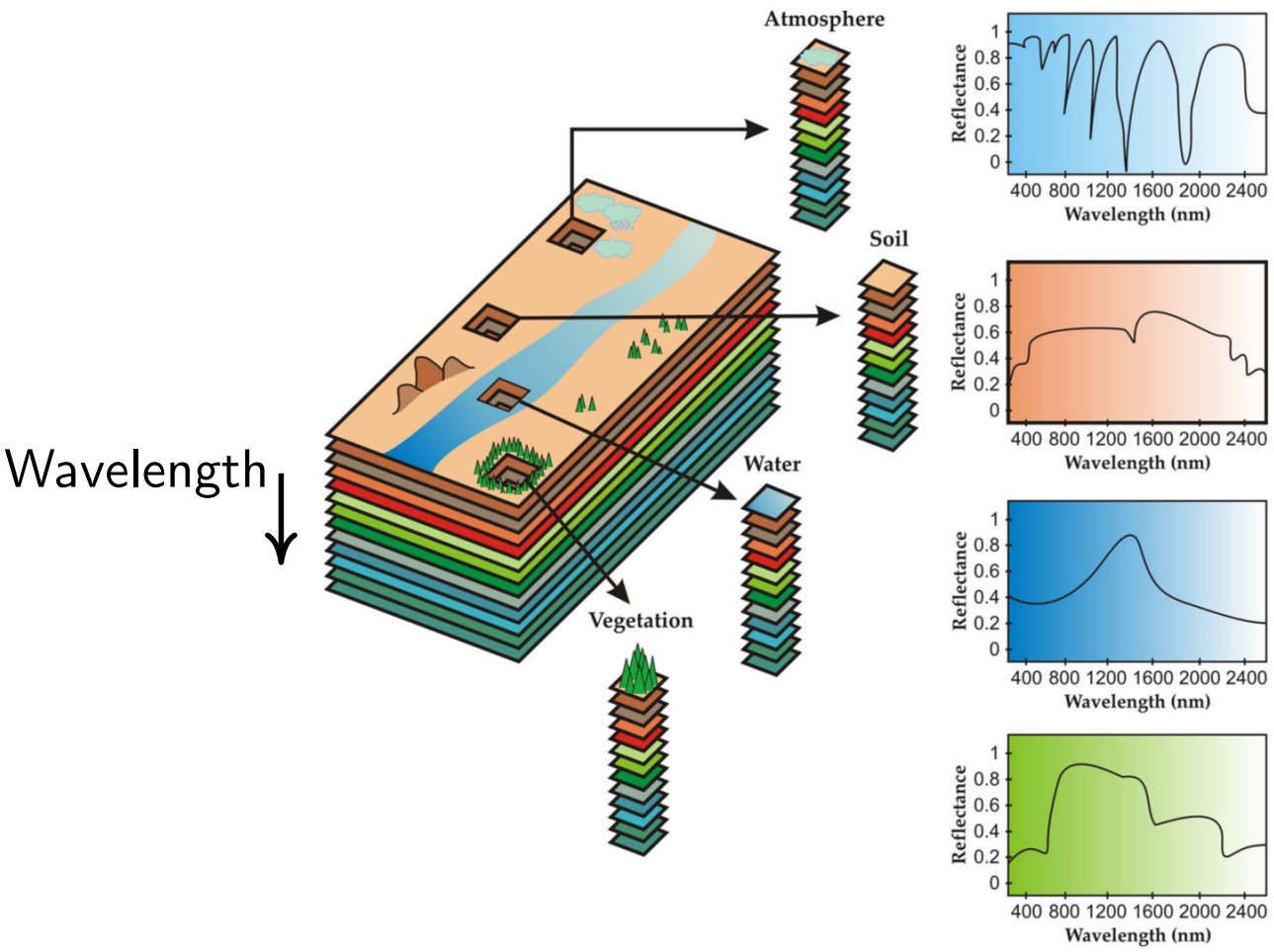


Figure: Concept of a spectral image [2].

Applications: Exploring Resources

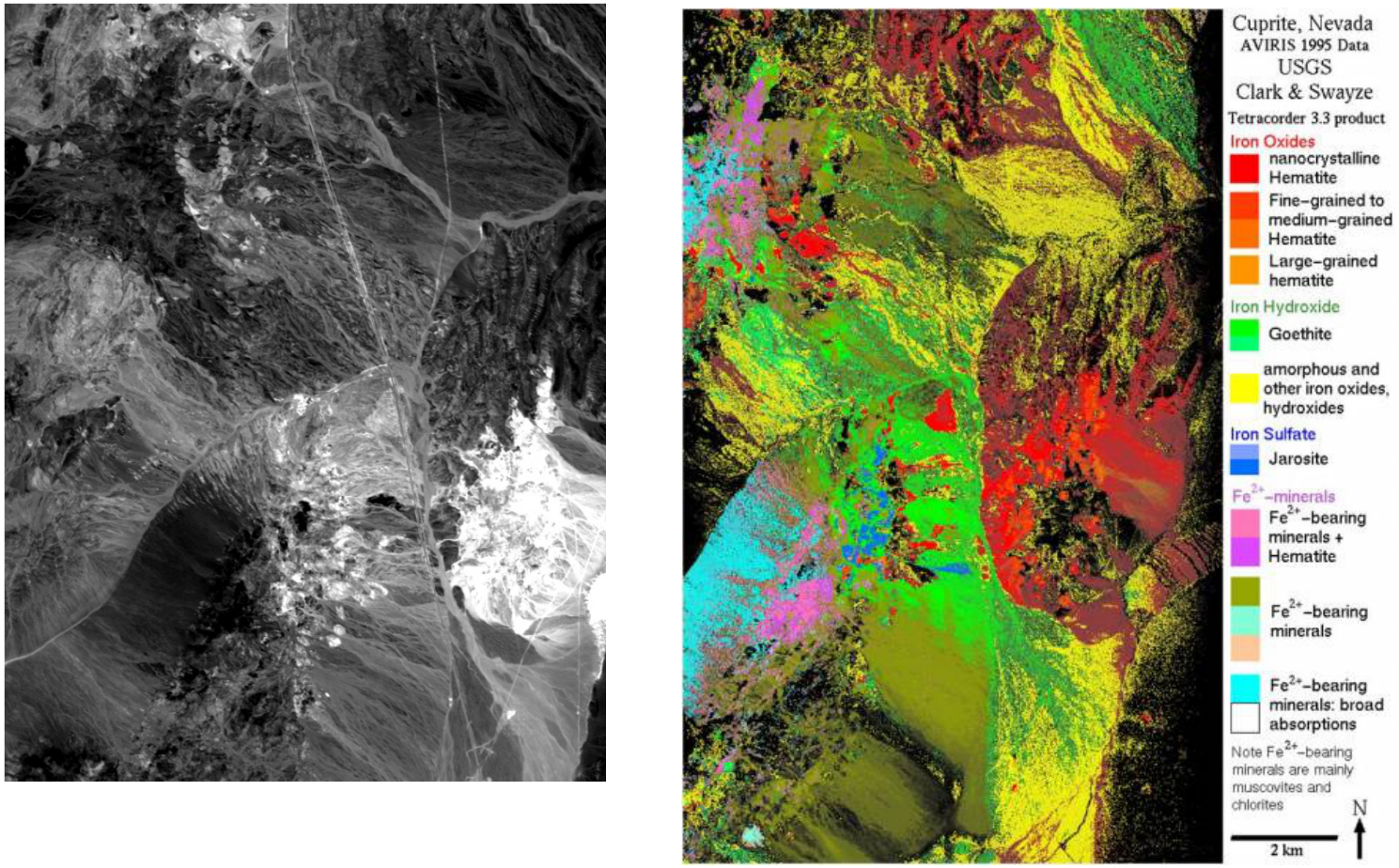
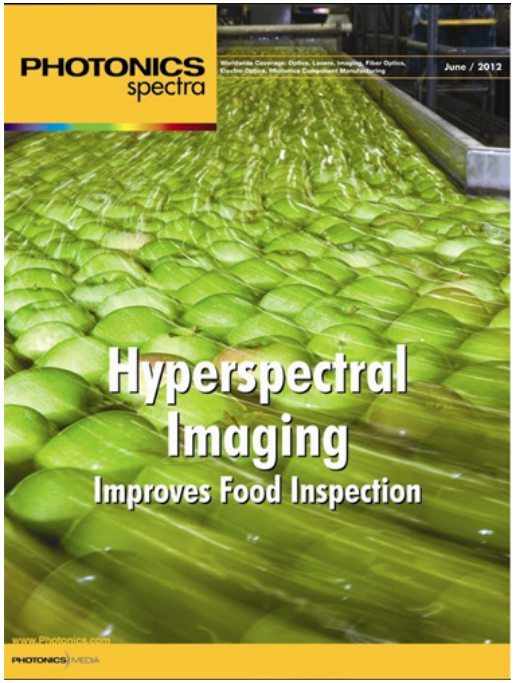


Figure: Image of Cuprite, Nevada captured in 1995.
Left: $0.815\,\mu$m band image. Right: identified mineral map [3, 4].

Applications: Food Quality Monitoring



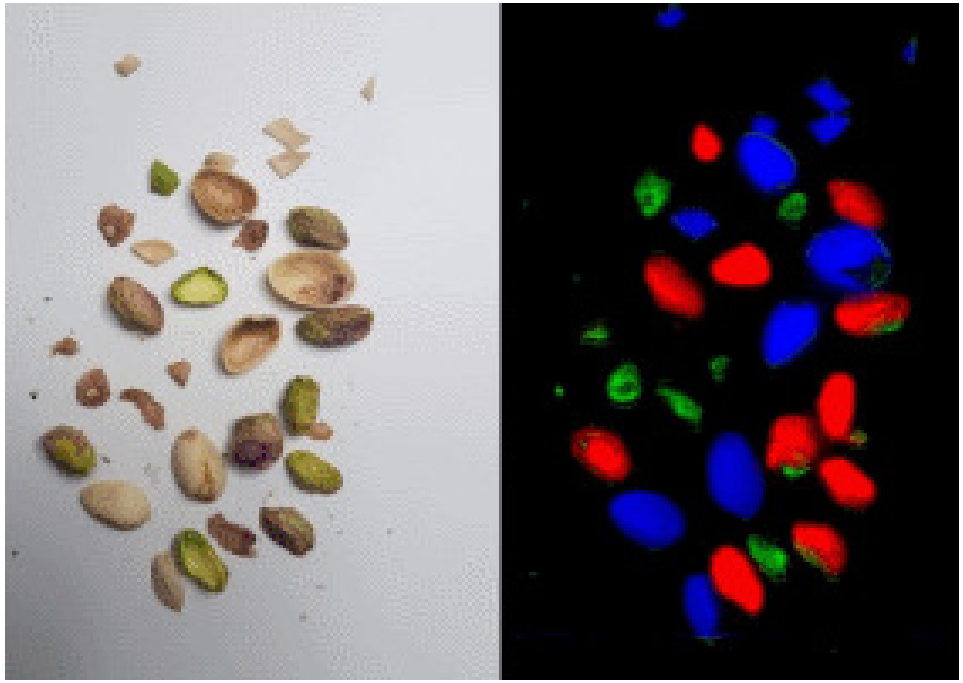


Figure: Food inspection, detection of foreign (unwanted) body.

Source: http://www.controlengeurope.com/article/132266/A-new-imaging-method-for-food-inspection.aspx

Applications: Wildfire Monitoring

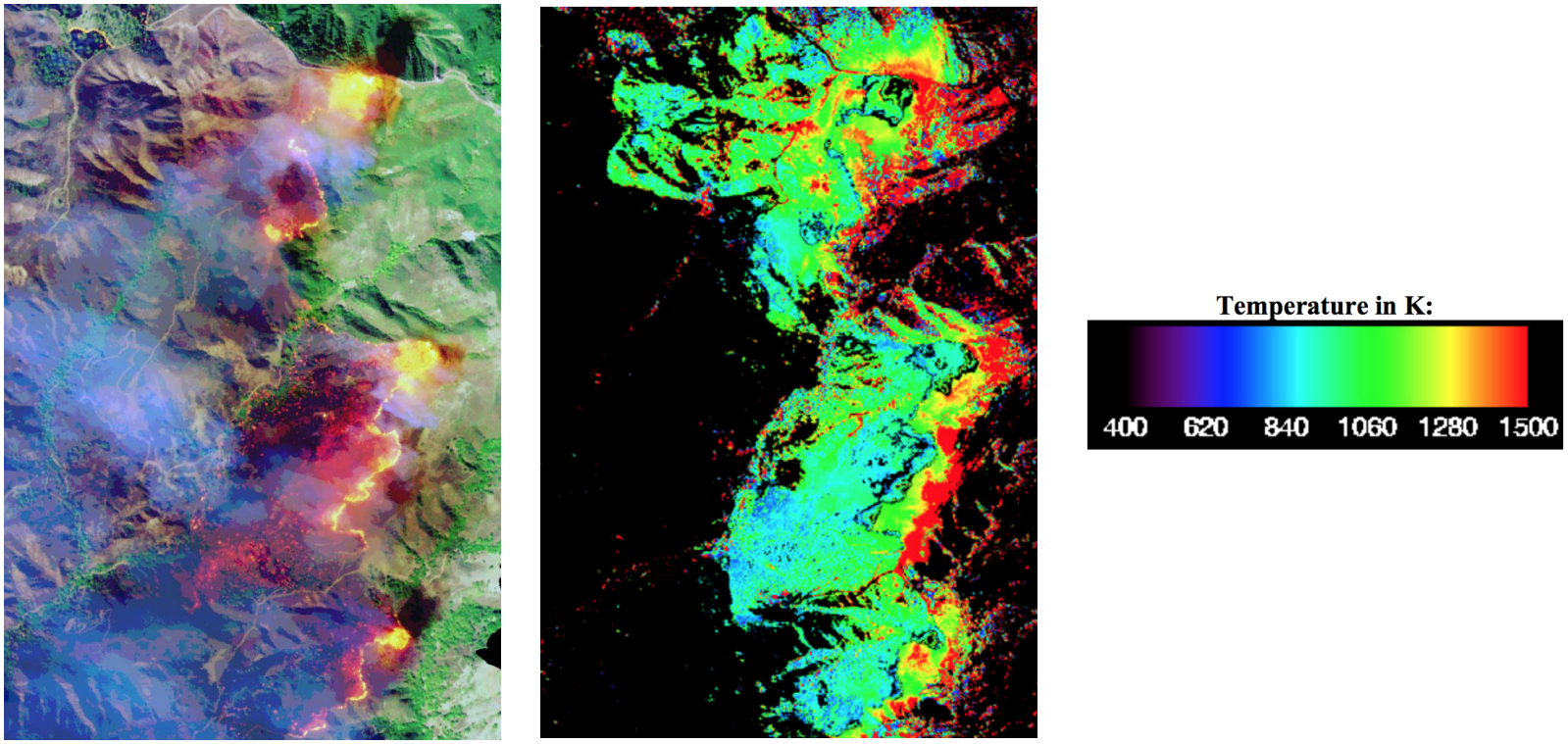
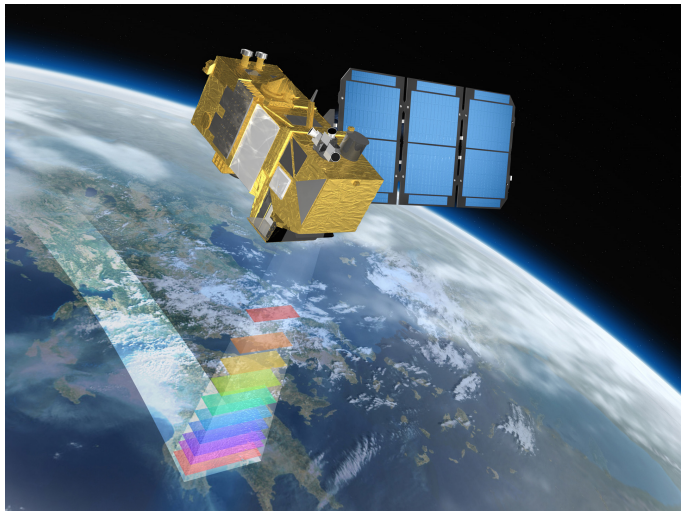


Figure: Identifying active fire fronts, post-fire hotspots and burnt areas [5]. A section of an AVIRIS scene acquired over the Simi Fire on Oct. 27, 2003. Temperature map is retrieved by using CO_2 signature to screen for actively burning pixels.

Spaceborne and Airborne Spectral Sensors

- an array of discrete sensors carried by a remote satellite;
- scan the earth surface along the flight path;



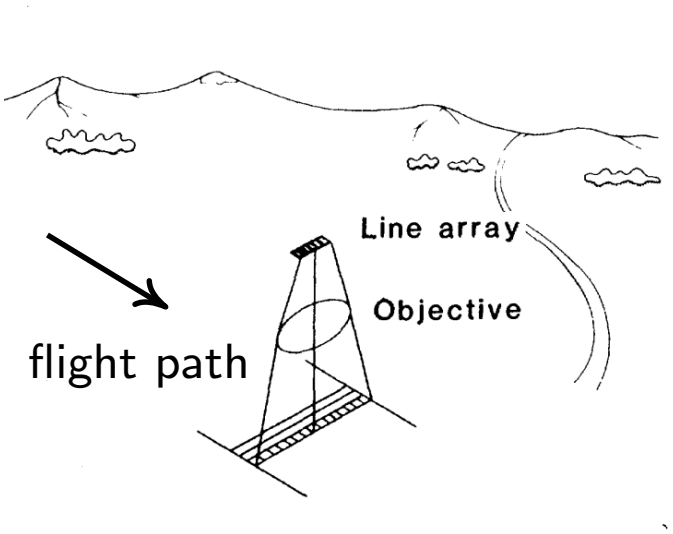


Figure: Line array sensors scanning the earth surface [6].

Spaceborne and Airborne Spectral Sensors

Spaceborne spectral sensors

- farther away from the ground, more spectral bands,
- called hyperspectral (HS) sensors,
- e.g.: Hyperion, ~ 700km above ground, > 200 bands,
  30m ground sampling distance (GSD);

Figure: Resolution difference between HS and MS images.

Spaceborne and Airborne Spectral Sensors

Airborne spectral sensors

- airborne: closer to the ground, less spectral bands.
- called multispectral (MS) sensors,
- e.g.: QuickBird, ~ 400km above ground, < 10 bands, 2.16m GSD.

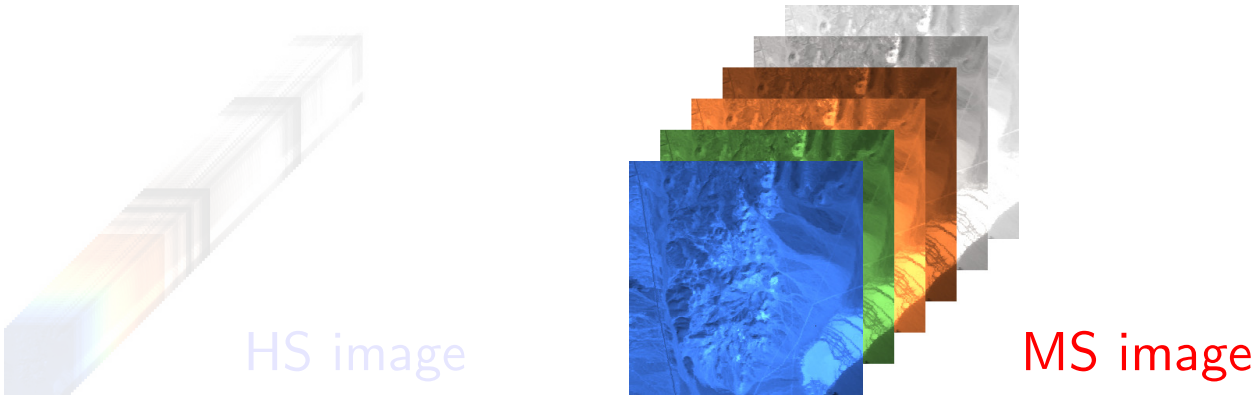


Figure: Resolution difference between HS and MS images.

Spaceborne and Airborne Spectral Sensors

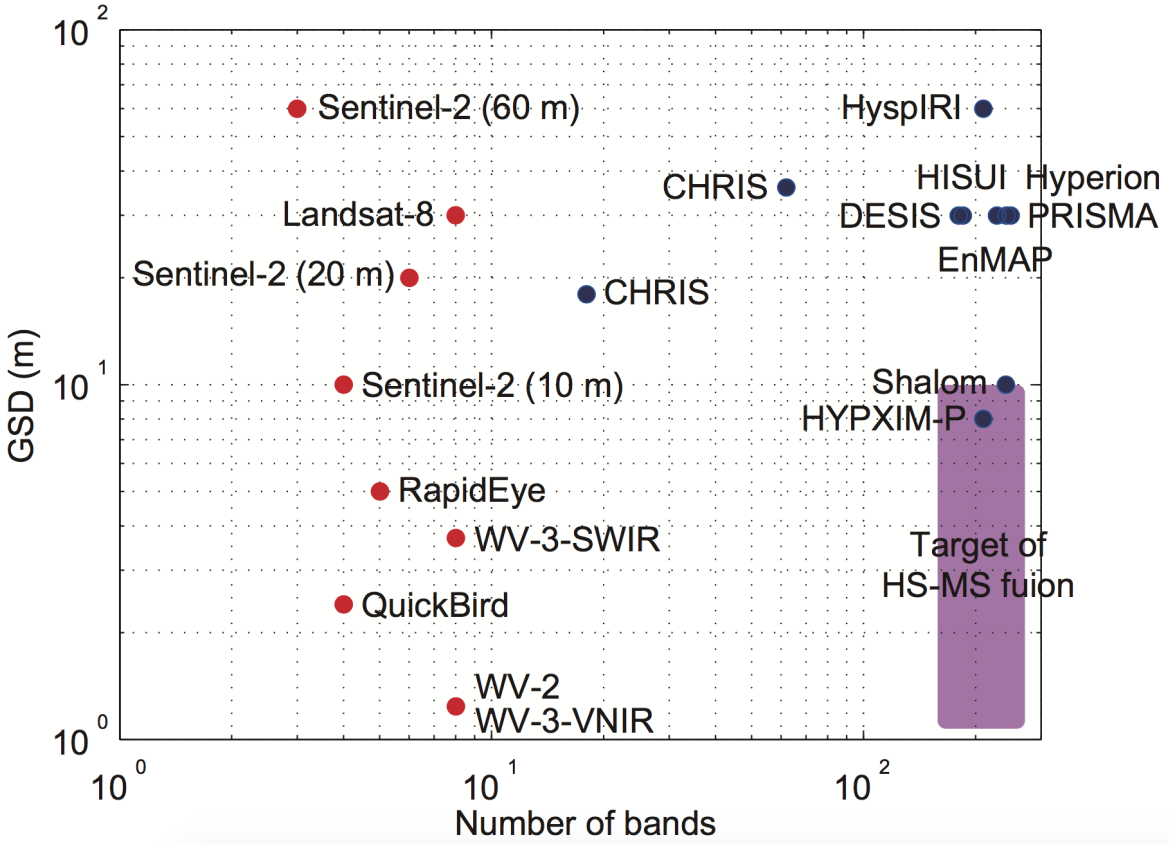


Figure: Sensor characteristics of sensors in latest spaceborne projects. GSD: Ground Sampling Distance [7].

Problem: High Resolution Image

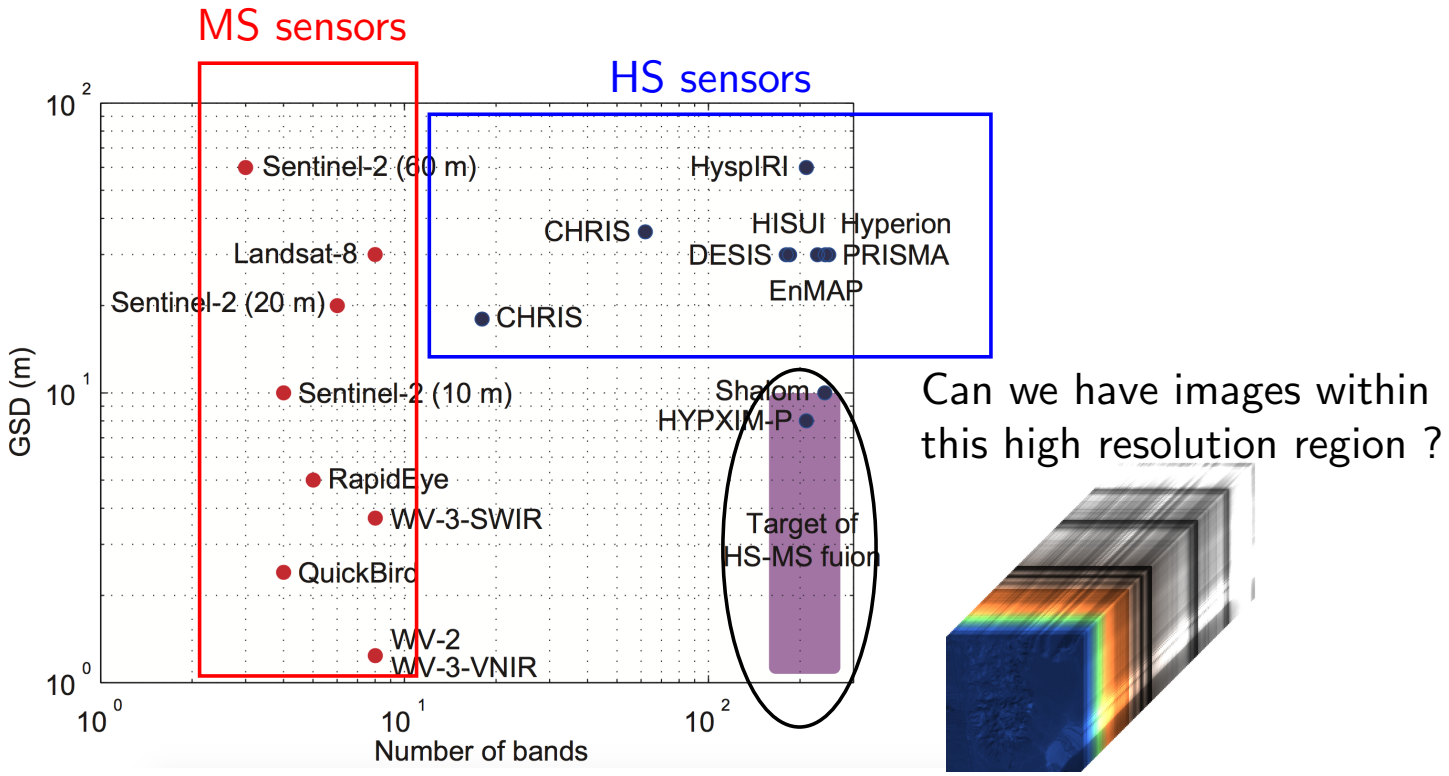


Figure: Sensor characteristics of sensors in latest spaceborne projects.
GSD: Ground Sampling Distance [7].

Outline

Introduction and Motivation
- Spectral Signature and Spectral Image
- Applications
- The Instruments
- What is the Problem ?

Hyperspectral Super-resolution (HSR)
- The HSR Problem
- State of the Arts

Fast Algorithm Framework of HSR (proposed)
- Preliminaries
- A Very Fast Algorithm Framework

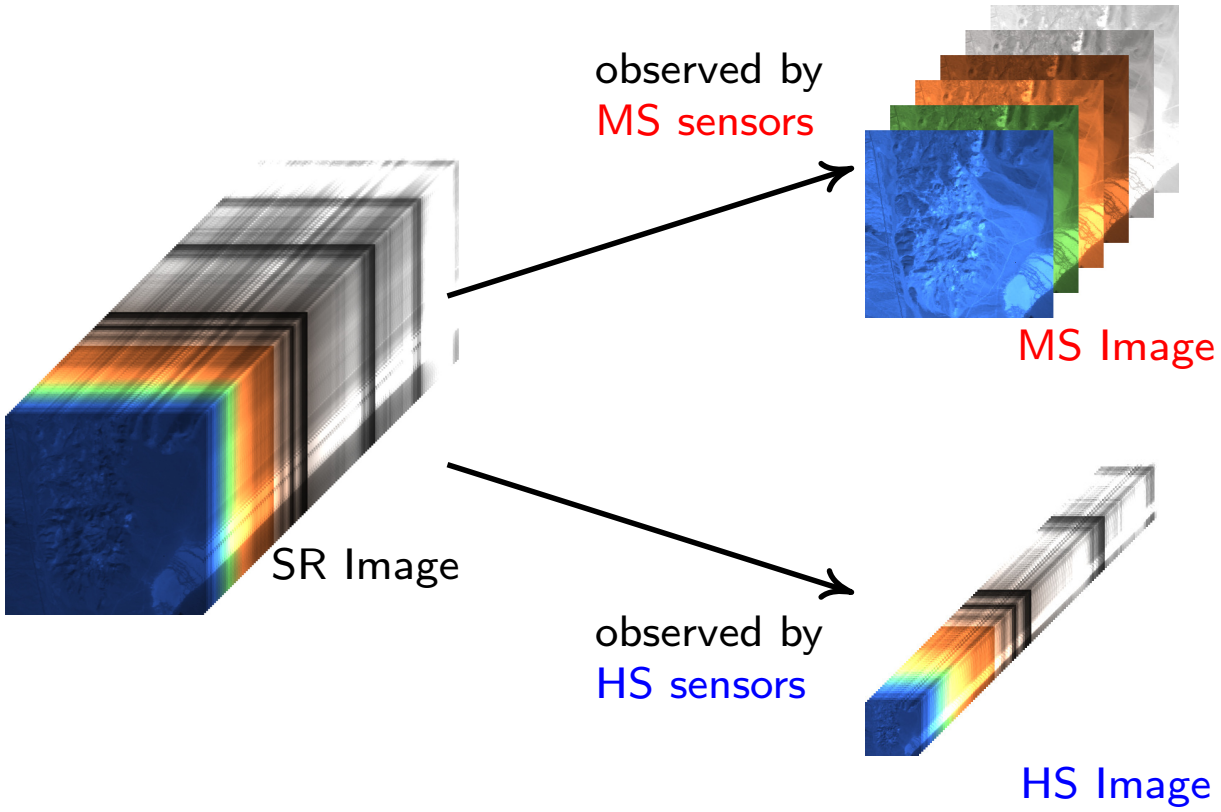
Experiments

Summary

Appendices and References

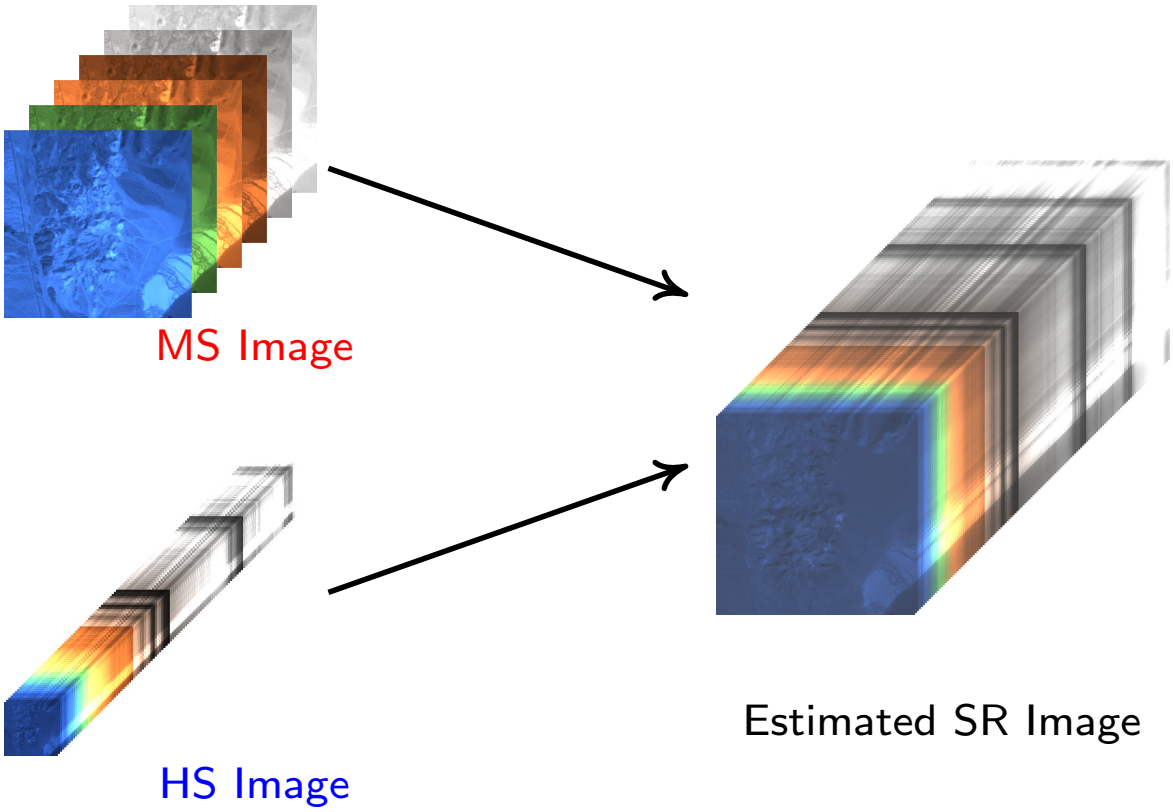
The HSR Problem

The hyperspectral super-resolution (HSR) problem is to estimate the super-resolution (SR) image from its HS image and MS image.



The HSR Problem

The hyperspectral super-resolution (HSR) problem is to estimate the super-resolution (SR) image from its HS image and MS image.



Signal and Model

Spectral image as tensor $\mathcal{Y} \in \mathbb{R}^{\ell_1 \times \ell_2 \times M}$, can be unfolded to $\boldsymbol{Y} \in \mathbb{R}^{M \times L}$. $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{y}_1 & \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_L \end{bmatrix}$ are the signatures from pixel 1 to pixel L.

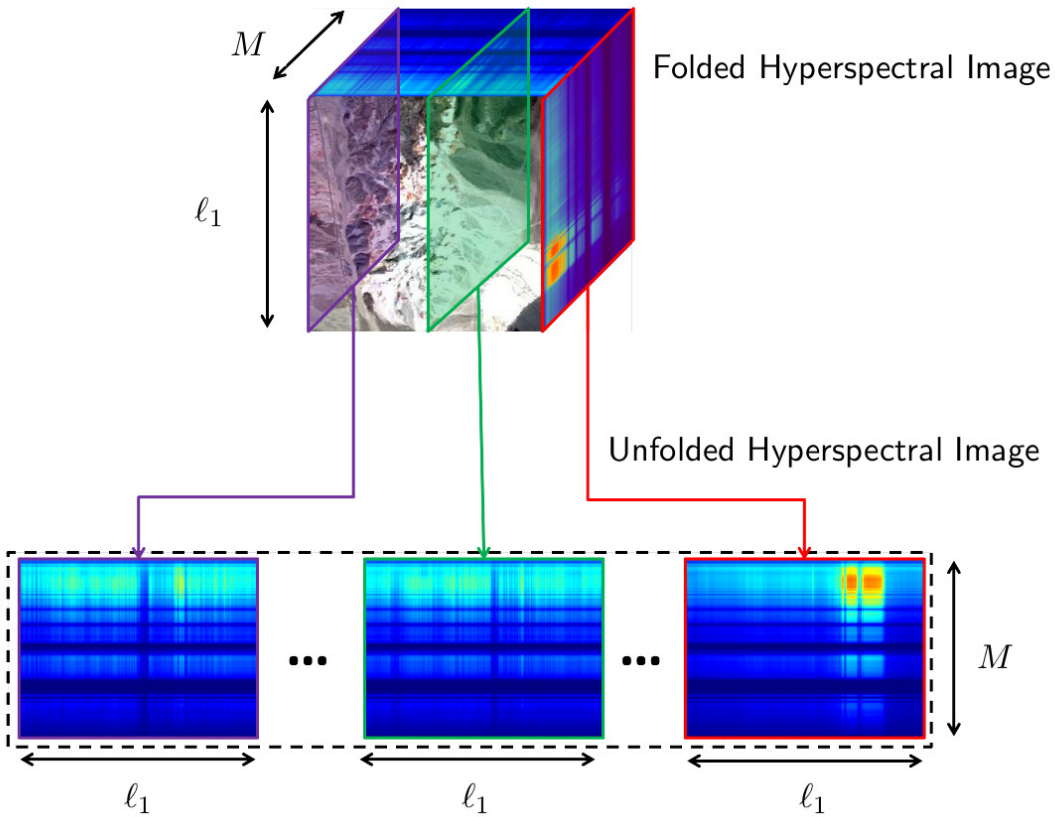


Figure: Unfolding tensor $\mathcal{Y}$ to matrix $\boldsymbol{Y}$. M as spectral band number; L as pixel number.

Signal and Model

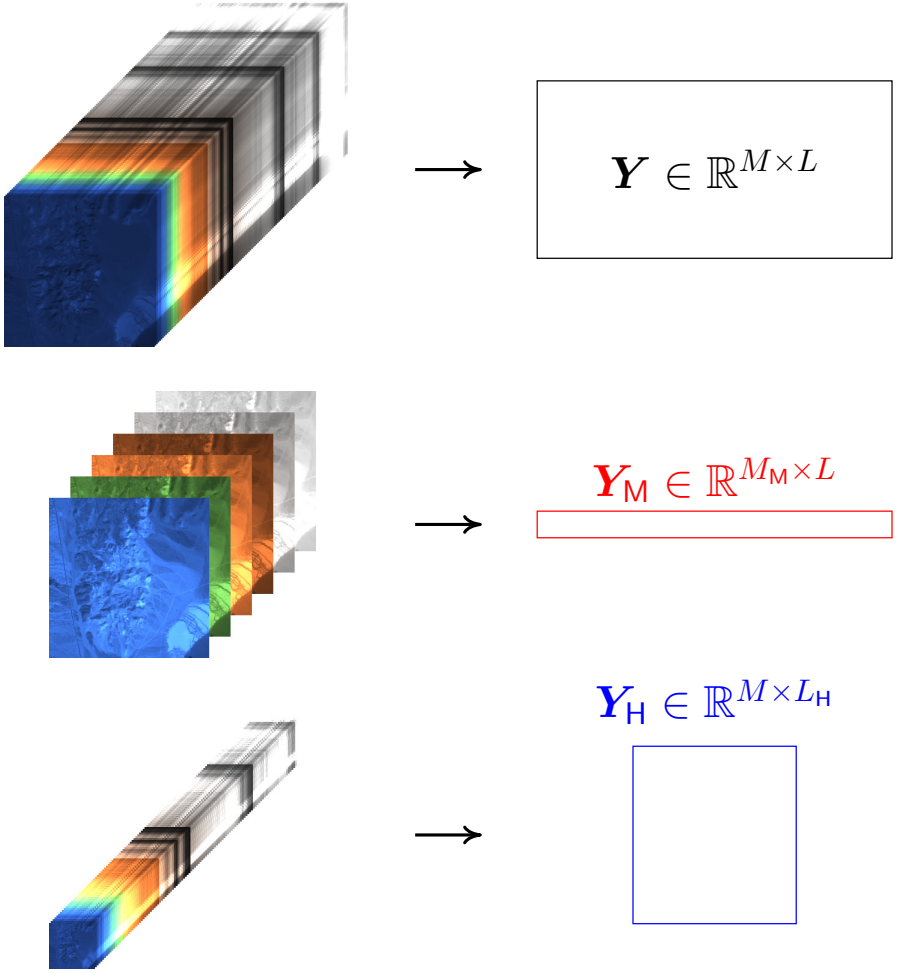


Figure: Unfolding HS and MS tensor to matrix with $M_{\mathsf{M}} < M$ and $L_{\mathsf{H}} < L$.

Signal and Model

We may relate $\boldsymbol{Y}$ and $\boldsymbol{Y}_{\mathsf{H}}$, $\boldsymbol{Y}_{\mathsf{M}}$ by

$$\boldsymbol{Y}_{\mathsf{H}} = \boldsymbol{Y}\boldsymbol{G},$$

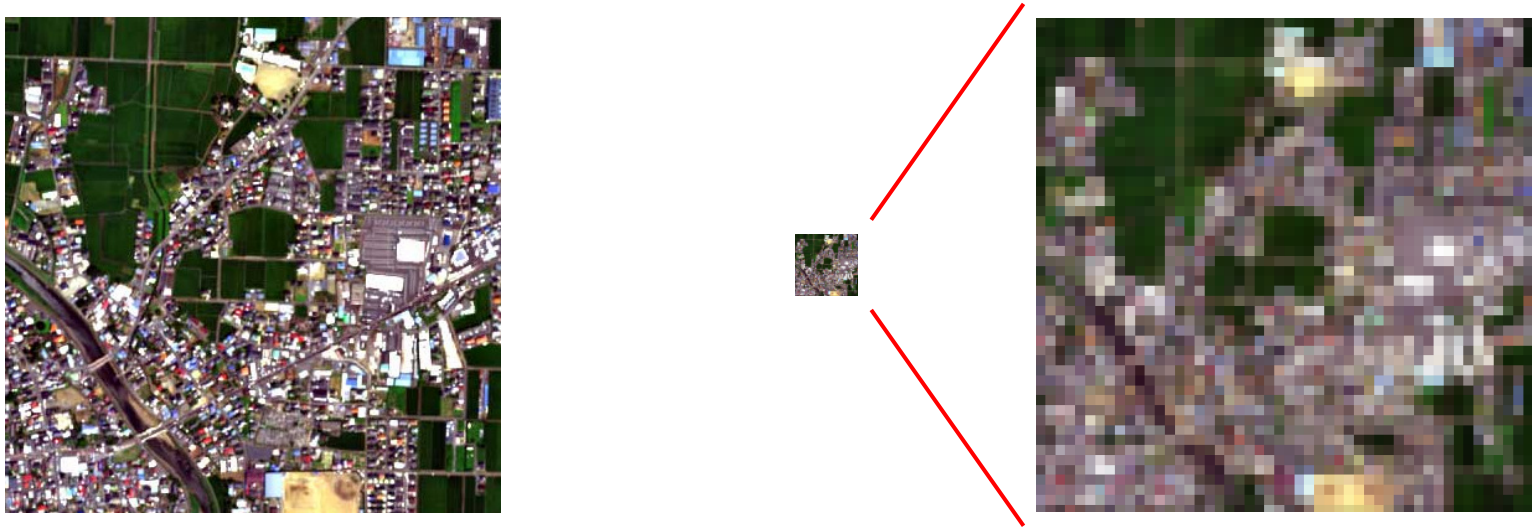
$$\boldsymbol{Y}_{\mathsf{M}} = \boldsymbol{F}\boldsymbol{Y},$$

where

- $\boldsymbol{G} \in \mathbb{R}^{L \times L_{\mathsf{H}}}$ is a tall matrix;
- multiplying with $\boldsymbol{G}$ serves as spatial downsampling;
- $\boldsymbol{G}$ simulates the response of an HS sensor.

Signal and Model

Spatial downsampling (blurring + subsampling) as HS sensor response.
Effect of $\boldsymbol{G}$ on an RGB satellite image:



(a)Full resolution image, 400×400 pixels. (b)Downsampled image, 50×50 pixels.

Figure: Downsampling effect.

Signal and Model

We may relate $\boldsymbol{Y}$ and $\boldsymbol{Y}_{\mathsf{H}}$, $\boldsymbol{Y}_{\mathsf{M}}$ by

$$\boldsymbol{Y}_{\mathsf{H}} = \boldsymbol{Y}\boldsymbol{G},$$

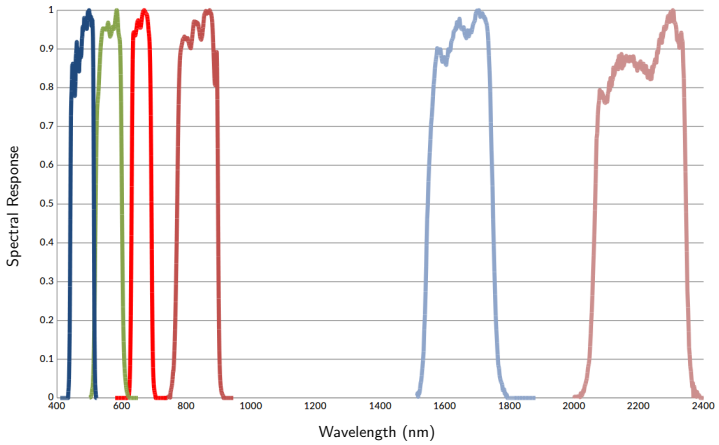
$$\boldsymbol{Y}_{\mathsf{M}} = \boldsymbol{F}\boldsymbol{Y},$$

where

- $\boldsymbol{F} \in \mathbb{R}^{M_{\mathsf{M}} \times M}$ is a fat matrix;
- being multiplied by $\boldsymbol{F}$ serves as spectral downsampling;
- $\boldsymbol{F}$ simulates the response of an MS sensor.

Signal and Model

Spectral downsampling (taking broad band image) as MS sensor response.



Wavelength (μm)	Spectral Band Placement
$0.45 - 0.52$	Visible (VIS-B)
$0.52 - 0.60$	Visible (VIS-G)
$0.63 - 0.69$	Visible (VIS-R)
$0.76 - 0.90$	Near Infrared (NIR)
$1.55 - 1.75$	Short Wave Infrared (SWIR-1)
$2.08 - 2.35$	Short Wave Infrared (SWIR-2)

Figure: Spectral response of the TM sensor in LANDSAT 7 Project [8, 9]. In this case there are 6 broad bands in the MS image.

Formulating the HSR Problem

Literatures have reviewed a number of HSR methods in [7, 10] including

- Component Substitution (CS), PCA-type Approaches
  $\rightarrow$ classical approach, often have spectral distortions;
- Multi-Resolution Analysis (MRA), Bayesian Approaches, Matrix Factorization (MF)
  $\rightarrow$ computationally expensive.

Among the above, Bayessian and MF approaches share similar problem structure $\rightarrow$ the focus of this thesis research.

Formulating the HSR Problem

A direct way of writing the HSR problem as an MF problem:

$$\min_{\boldsymbol{Y}} \quad \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{Y}\boldsymbol{G}\|_{\mathsf{F}}^2 + \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{Y}\|_{\mathsf{F}}^2,$$

which is a convex but *underdetermined* problem.

We can get around this issue by making assumptions on $\boldsymbol{Y}$.

Formulating the HSR Problem

In general, we know that $\boldsymbol{Y} \geq \boldsymbol{0}$ because light intensity is nonnegative.

$\exists$ strong belief of the Linear Mixing Model (LMM) of $\boldsymbol{Y}$ in literatures, i.e.,

- $\boldsymbol{Y}$ has low-rank;
- $\{\, \boldsymbol{y}[\, j\,] \,\}_{j=1}^{L}$ are linear mixtures of $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N\}$, with $N \ll \min(M, L)$.

$$\boldsymbol{y}[\, j\,] = \sum_{i=1}^{N} \boldsymbol{a}_i \; s_i[\, j\,] \quad \forall\, j = 1, \cdots, L.$$

Formulating the HSR Problem

LMM Interpretation:

- $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N\}$ as "endmember" signatures, the spectral signature of pure material;
- measurement in each pixels is a mixture of pure materials signature.

$$
\begin{aligned}
\boldsymbol{y}[\,j\,] &= \sum_{i=1}^{N} \boldsymbol{a}_i \; s_i[\,j\,], \\
\boldsymbol{Y} &= \begin{bmatrix} \boldsymbol{y}[\,1\,] & \cdots & \boldsymbol{y}[\,L\,] \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_N \end{bmatrix}}_{\text{endmember}} \underbrace{\begin{bmatrix} \boldsymbol{s}[\,1\,] & \cdots & \boldsymbol{s}[\,L\,] \end{bmatrix}}_{\text{abundance}} = \boldsymbol{A}\,\boldsymbol{S}.
\end{aligned}
$$

Formulating the HSR Problem

LMM Interpretation:

- $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N\}$ as "endmember" signatures, the spectral signature of pure material;
- measurement in each pixels is a mixture of pure materials signature.

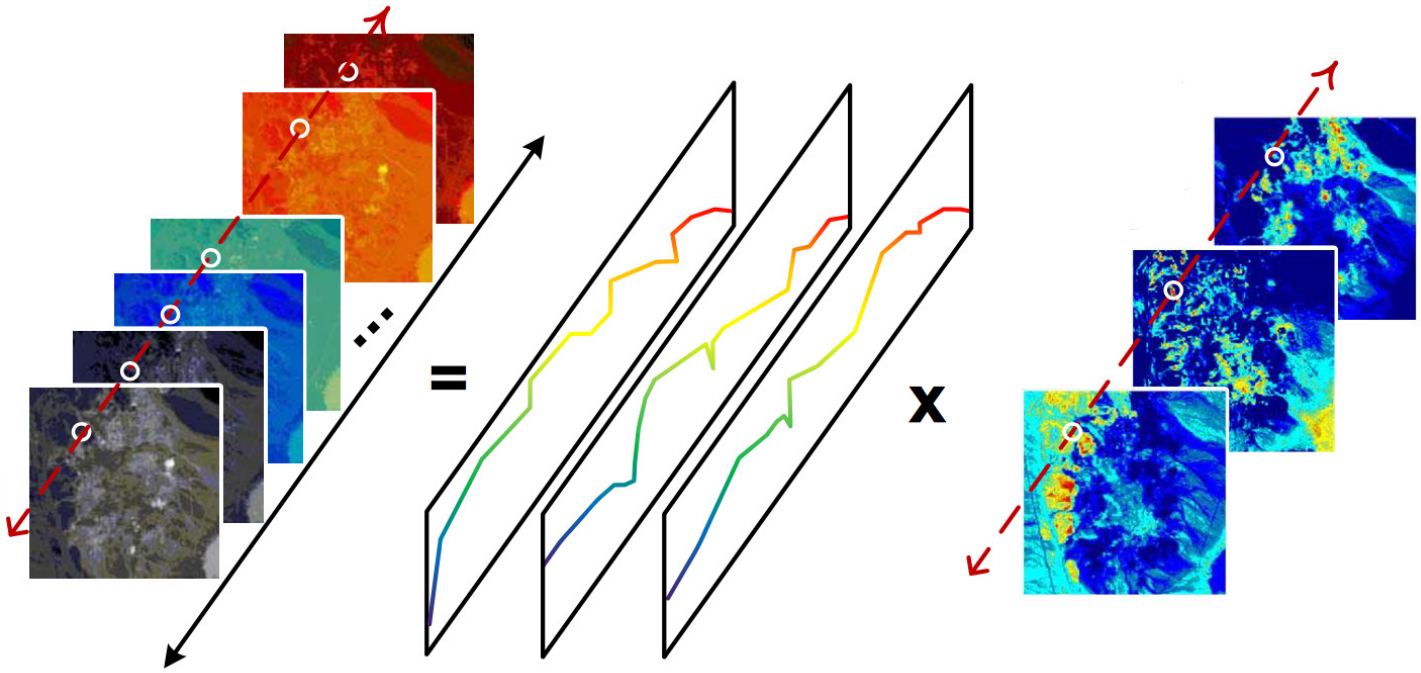


Figure: Linear Mixing Model. [11]

Formulating the HSR Problem

Is LMM true ?

- not necessarily true;
- affected by ground surface material alignment.
- many literatures believe so, have successful demonstration of analysis in laboratory-based experiments.

$$\boldsymbol{Y} \quad = \quad \boldsymbol{A}\,\boldsymbol{S}.$$

Formulating the HSR Problem

Employing the LMM changes the HSR problem from

$$\boxed{\min_{\boldsymbol{Y}} \quad \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{Y}\boldsymbol{G}\|_{\mathsf{F}}^2 + \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{Y}\|_{\mathsf{F}}^2. \quad \text{(convex, ill-posed)}}$$

to

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{A},\boldsymbol{S}} & \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2 \quad \text{(nonconvex, less ill-posed)} \\ \text{s.t.} & \boldsymbol{A} \in \mathcal{A},\ \boldsymbol{S} \in \mathcal{S}. \end{array}}$$

Formulating the HSR Problem

$$\min_{\boldsymbol{A},\boldsymbol{S}} \ \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{ASG}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{FAS}\|_{\mathsf{F}}^2$$
$$\text{s.t.} \ \ \boldsymbol{A} \in \mathcal{A}, \ \boldsymbol{S} \in \mathcal{S}.$$

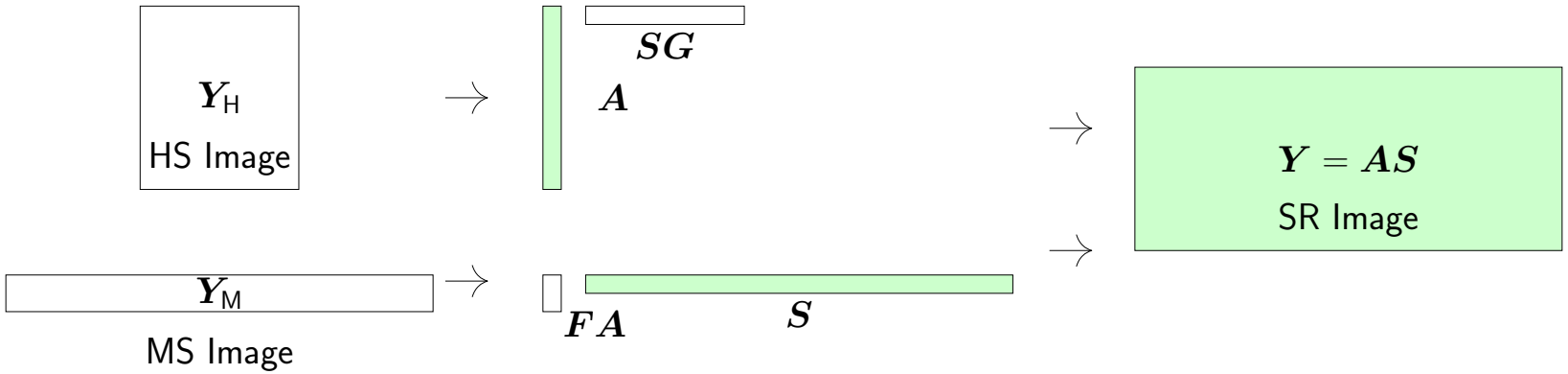


Figure: Estimate SR image from HS and MS image pair.

State of the Arts HSR

Two state-of-the-art HSR algorithms: find $\hat{\boldsymbol{Y}} = \boldsymbol{A}\boldsymbol{S}$.

- Coupled Nonnegative Matrix Factorization (CNMF) [12], solves

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{A},\boldsymbol{S}} & \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2 \\ \text{s.t.} & \boldsymbol{A} \geq \boldsymbol{0},\ \boldsymbol{S} \geq \boldsymbol{0}; \end{array}}$$

- Fusion based on spectral UnMIxing (FUMI) [13], solves

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{A},\boldsymbol{S}} & \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2 \\ \text{s.t.} & \boldsymbol{0} \leq \boldsymbol{A} \leq \boldsymbol{1},\ \boldsymbol{S} \geq \boldsymbol{0},\ \boldsymbol{1}^\top \boldsymbol{S} = \boldsymbol{1}^\top. \end{array}}$$

State of the Arts HSR

CNMF algorithm [12]: find $\hat{\boldsymbol{Y}} = \boldsymbol{A}\boldsymbol{S}$.

- does not strictly solve the HSR problem;
- splits HSR problem into two NMF problems, i.e., alternatingly solves

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{A},\boldsymbol{R}} & \|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{R}\|_{\mathsf{F}}^2 \\ \text{s.t.} & \boldsymbol{A}, \boldsymbol{R} \geq \boldsymbol{0} \end{array}} \rightleftharpoons \boxed{\begin{array}{ll} \min\limits_{\boldsymbol{B},\boldsymbol{S}} & \|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{B}\boldsymbol{S}\|_{\mathsf{F}}^2 \\ \text{s.t.} & \boldsymbol{B}, \boldsymbol{S} \geq \boldsymbol{0} \end{array}}$$

- an alternating optimization (AO) on 2 different problems.
- pros: a fast method available: Lee-Seung Multiplicative Update [14]
  details
- cons: a heuristic apprach, no convergence guarantee from optimization view point.

State of the Arts HSR

FUMI algorithm [13]: find $\hat{\boldsymbol{Y}} = \boldsymbol{A}\boldsymbol{S}$.

- alternatingly solves

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{A}} & \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2 \\ \text{s.t.} & \boldsymbol{0} \leq \boldsymbol{A} \leq \boldsymbol{1}, \end{array}}$$

$$\Updownarrow$$

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{S}} & \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2 \\ \text{s.t.} & \boldsymbol{S} \geq \boldsymbol{0},\ \boldsymbol{1}^\top \boldsymbol{S} = \boldsymbol{1}^\top \end{array}}$$

- an AO on the same problem;
- uses ADMM to solve;
- pros: has convergence guarantee.
- cons: unsuitable for big matrix problems ($\boldsymbol{G}$ is potentially large);
- cons: highly depends on special structure of $\boldsymbol{G}$.

State of the Arts HSR

Some more state of the arts ...

Sparse Promoting NMF [15]:

$$\begin{aligned} \min_{\boldsymbol{A},\boldsymbol{S}} \quad & \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2 + \lambda\, \mathsf{sparse}(\boldsymbol{S}) \\ \text{s.t.} \quad & \boldsymbol{A},\ \boldsymbol{S} \geq \boldsymbol{0}; \end{aligned}$$

HYperspectral SUper-REsolution (HYSURE) [16]:
with an estimated endmember $\hat{\boldsymbol{A}}$,

$$\min_{\boldsymbol{S}} \quad \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \hat{\boldsymbol{A}}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \tfrac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\hat{\boldsymbol{A}}\boldsymbol{S}\|_{\mathsf{F}}^2 + \lambda\, \mathsf{TV}(\boldsymbol{S}).$$

Outline

Introduction and Motivation
- Spectral Signature and Spectral Image
- Applications
- The Instruments
- What is the Problem ?

Hyperspectral Super-resolution (HSR)
- The HSR Problem
- State of the Arts

Fast Algorithm Framework of HSR (proposed)
- Preliminaries
- A Very Fast Algorithm Framework

Experiments

Summary

Appendices and References

Preliminaries

Let $f(\boldsymbol{A}, \boldsymbol{S}) = \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2$ and $\mathcal{S}$, $\mathcal{A}$ are convex.

We focus on the basic HSR problem:

$$\begin{aligned} \min_{\boldsymbol{A},\boldsymbol{S}} \quad & f(\boldsymbol{A}, \boldsymbol{S}) \\ \text{s.t.} \quad & \boldsymbol{A} \in \mathcal{A},\ \boldsymbol{S} \in \mathcal{S}. \end{aligned} \tag{1}$$

Traditionally, people use AO to solve (1), i.e., iterate for $k = 0, 1, \cdots$

$$\boldsymbol{S}^{(k+1)} \leftarrow \min_{\boldsymbol{S} \in \mathcal{S}} \quad f(\boldsymbol{A}^{(k)}, \boldsymbol{S}) \tag{2}$$

$$\boldsymbol{A}^{(k+1)} \leftarrow \min_{\boldsymbol{A} \in \mathcal{A}} \quad f(\boldsymbol{A}, \boldsymbol{S}^{(k+1)}) \tag{3}$$

Iterating (2) and (3) can be seen as a 2-block coordinate descent (BCD).

Preliminaries

HSR by BCD on (2) and (3):

- successive minimization along direction of $\boldsymbol{S}$ and $\boldsymbol{A}$;
- $f(\boldsymbol{A}, \boldsymbol{S})$ is individually quadratic in $\boldsymbol{A}$ or in $\boldsymbol{S}$;
- global solutions to (2) and (3) are available ($\because$ they are QP ▸ details).

Pros:

- can exactly solve for global solutions *if (2),(3) are small*;
- can employ ADMM, gradient methods or higher order methods (e.g. Newton's method, interior point methods) *if (2),(3) are small*.

Cons:

- big data era, interested in big spectral image processing;
- **hard** to exactly solve for global solutions *when (2), (3) are big*;
- higher order methods are **unsuitable** *when (2), (3) are big*.

Fast Framework of HSR (proposed)

Fast framework:

- HSR problem as AO on two QPs, solved by *Inexact* BCD;
- uses 1st order gradient methods, which can be
  - Gradient Projection (GP) type methods:
    - GP [17],
    - Barzilai-Borwein GP (BBGP) [18];
  - Proximal Gradient (PG) type methods:
    - PG [19],
    - Fast PG (FISTA) [20, 21];
  - Projection-free type method:
    - Frank-Wolfe (FW) [22].

Fast Framework of HSR (proposed)

A brief overview on 1st order gradient methods.
Suppose $g(\boldsymbol{x})$ is quadratic in $\boldsymbol{x}$, and $\mathcal{X}$ is the feasible set of $\boldsymbol{x}$.
We repeat, for $k = 0, 1, \cdots$,

$$\boldsymbol{x}^{(k+1)} \leftarrow \boldsymbol{x}^{(k)} + \gamma^{(k)} \boldsymbol{d}^{(k)}, \quad \text{(GP, BBGP and FW, with } \boldsymbol{d}^{(k)} \in \mathcal{X}) \tag{4}$$

or

$$\boldsymbol{x}^{(k+1)} \leftarrow \Pi_{\mathcal{X}} \left(\boldsymbol{x}^{(k)} + \gamma^{(k)} \boldsymbol{d}^{(k)} \right). \quad \text{(PG and FISTA, with } \boldsymbol{d}^{(k)} \in \text{ or } \notin \mathcal{X}) \tag{5}$$

- $\boldsymbol{d}^{(k)}$ is the descent direction, can be the direction opposite to $\nabla g(\boldsymbol{x}^{(k)})$, or any directions that guarantee $g(\boldsymbol{x}^{(k+1)}) \leq g(\boldsymbol{x}^{(k)})$;
- $\gamma^{(k)}$ is the step size;

Fast Framework of HSR (proposed)

A brief overview on 1st order gradient methods.

1st order method	GP	BBGP	PG	FISTA	FW
requires $\nabla g(\boldsymbol{x}^{(k)})$ computation	✓	✓	✓	✓	✓
requires projection in finding $\boldsymbol{d}^{(k)}$	✓	✓	✗	✗	✗
requires projection in finding $\boldsymbol{x}^{(k+1)}$	✗	✗	✓	✓	✗
requires searching for $\gamma^{(k)}$	✓	✓	✗	✗	✓

Fast Framework of HSR (proposed)

Traditionally, people do the following:

Algorithm 1 HSR by Exact BCD with 1st order gradient methods

procedure
Input: $\boldsymbol{S}^{(0)} \in \mathcal{S}$, $\boldsymbol{A}^{(0)} \in \mathcal{A}$; Output: $\boldsymbol{A}^{(K+1)}$, $\boldsymbol{S}^{(K+1)}$.
  for $k = 0, 1, 2, \cdots, K$ **do**
    for $j = 1, 2, \cdots, J$ **do**
      $\boldsymbol{S}^{(k)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k)}))$.
    end for
    $\boldsymbol{S}^{(k+1)} \leftarrow \boldsymbol{S}^{(k)}$.
    for $j = 1, 2, \cdots, J$ **do**
      $\boldsymbol{A}^{(k)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k+1)}))$.
    end for
    $\boldsymbol{A}^{(k+1)} \leftarrow \boldsymbol{A}^{(k)}$.
  end for
end procedure

Fast Framework of HSR (proposed)

The proposed fast framework does the following: HSR by Inexact BCD with 1st order methods (named ALternating Gradient-based Optimization ALGO)

Algorithm 2 HSR by ALGO

procedure
Input: $\boldsymbol{S}^{(0)} \in \mathcal{S}$, $\boldsymbol{A}^{(0)} \in \mathcal{A}$; Output: $\boldsymbol{A}^{(K+1)}$, $\boldsymbol{S}^{(K+1)}$.
  for $k = 0, 1, 2, \cdots, K$ **do**
    $\boldsymbol{S}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k)}))$.
    $\boldsymbol{A}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k+1)}))$.
  end for
end procedure

Fast Framework of HSR (proposed)

HSR by ALGO

procedure
Input: $\boldsymbol{S}^{(0)} \in \mathcal{S}$, $\boldsymbol{A}^{(0)} \in \mathcal{A}$; Output: $\boldsymbol{A}^{(K+1)}$, $\boldsymbol{S}^{(K+1)}$.
  for $k = 0, 1, 2, \cdots, K$ **do**
    $\boldsymbol{S}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k)}))$.
    $\boldsymbol{A}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k+1)}))$.
  end for
end procedure

Remark:

- special case of Algorithm 1 when $J = 1$;
- `1st_GradUpd` is a **one-step** gradient update either in (4) or (5);
- in 1 outer iteration, ALGO is **much faster** than Exact BCD.

Fast Framework of HSR (proposed)

HSR by ALGO

procedure
Input: $\boldsymbol{S}^{(0)} \in \mathcal{S}$, $\boldsymbol{A}^{(0)} \in \mathcal{A}$; Output: $\boldsymbol{A}^{(K+1)}$, $\boldsymbol{S}^{(K+1)}$.
  for $k = 0, 1, 2, \cdots, K$ **do**
    $\boldsymbol{S}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k)}))$.
    $\boldsymbol{A}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k+1)}))$.
  end for
end procedure

Remark:

- with $f(\boldsymbol{A}, \boldsymbol{S}) = \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{H}} - \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\|_{\mathsf{F}}^2 + \frac{1}{2}\|\boldsymbol{Y}_{\mathsf{M}} - \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2$,
  gradient is just some algebrae:
  - $\nabla_{\boldsymbol{S}} f(\boldsymbol{A}, \boldsymbol{S}) = \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}\boldsymbol{G}^\top + (\boldsymbol{F}\boldsymbol{A})^\top \boldsymbol{F}\boldsymbol{A}\boldsymbol{S} - \boldsymbol{A}^\top \boldsymbol{Y}_{\mathsf{H}}\boldsymbol{G}^\top - (\boldsymbol{F}\boldsymbol{A})^\top \boldsymbol{Y}_{\mathsf{M}}$,
  - $\nabla_{\boldsymbol{A}} f(\boldsymbol{A}, \boldsymbol{S}) = \boldsymbol{A}\boldsymbol{S}\boldsymbol{G}(\boldsymbol{S}\boldsymbol{G})^\top + \boldsymbol{F}^\top \boldsymbol{F}\boldsymbol{A}\boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Y}_{\mathsf{H}}(\boldsymbol{S}\boldsymbol{G})^\top - \boldsymbol{F}^\top \boldsymbol{Y}_{\mathsf{M}}\boldsymbol{S}^\top$.

Fast Framework of HSR (proposed)

HSR by ALGO

procedure
Input: $\boldsymbol{S}^{(0)} \in \mathcal{S}$, $\boldsymbol{A}^{(0)} \in \mathcal{A}$; Output: $\boldsymbol{A}^{(K+1)}$, $\boldsymbol{S}^{(K+1)}$.
  for $k = 0, 1, 2, \cdots, K$ **do**
    $\boldsymbol{S}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k)}))$.
    $\boldsymbol{A}^{(k+1)} \leftarrow \texttt{1st_GradUpd}(f(\boldsymbol{A}^{(k)}, \boldsymbol{S}^{(k+1)}))$.
  end for
end procedure

Remark:

- `1st_GradUpd()` can be chosen as one of the following:
  - Grad. Proj. type: `GP_Upd()`, `BBGP_Upd()`;
  - Prox. Grad. type: `PG_Upd()`, `FISTA_Upd()`;
  - Proj. Free type: `FW_Upd()`.

Outline

Introduction and Motivation
- Spectral Signature and Spectral Image
- Applications
- The Instruments
- What is the Problem ?

Hyperspectral Super-resolution (HSR)
- The HSR Problem
- State of the Arts

Fast Algorithm Framework of HSR (proposed)
- Preliminaries
- A Very Fast Algorithm Framework

Experiments

Summary

Appendices and References

Experiments

HSR experimental objectives: to see

1) if we use 1st order gradient algorithms, which one is better ? Exact BCD or ALGO ?
2) comparing with CNMF & FUMI, can traditional BCD or ALGO have better performance ?

Performance measure:

- global measure: $\mathrm{RMSE}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) := \dfrac{\|\boldsymbol{Y} - \hat{\boldsymbol{Y}}\|_{\mathsf{F}}}{\sqrt{ML}}$;

  $\rightarrow$ lower the better
- spectral measure: $\mathrm{SAM}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) := \dfrac{1}{L} \displaystyle\sum_{j=1}^{L} \cos^{-1}\left(\dfrac{\langle \boldsymbol{y}[\, j\,] \,,\, \hat{\boldsymbol{y}}[\, j\,] \rangle}{\|\boldsymbol{y}[\, j\,]\|_2 \;\; \|\hat{\boldsymbol{y}}[\, j\,]\|_2} \right)$;

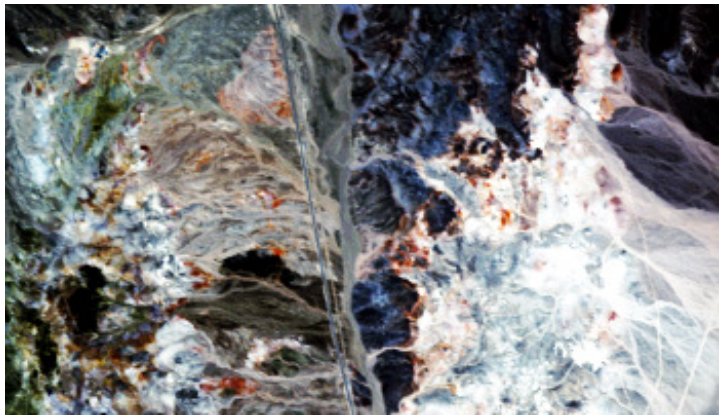
  $\rightarrow$ lower the better
- spatial measure: $\mathrm{PSNR}(\boldsymbol{Y}, \hat{\boldsymbol{Y}}) := \dfrac{1}{M} \displaystyle\sum_{i=1}^{M} 10 \log_{10} \left(\dfrac{\max(\boldsymbol{y}^i)}{\|\boldsymbol{y}^i - \hat{\boldsymbol{y}}^i\|_2^2 / L} \right)$.

  $\rightarrow$ higher the better

Experiments

Testing subjects:
from 2 real spectral image data sets: "Chikusei" [23] and "Cuprite" [4].

(a) Chikusei data set, with 128 spectral bands, 1000^2 pixels.



(b) Cuprite data set, with 224 spectral bands, 200×348 pixels.

Figure: Color image of Chikusei and Cuprite data set.

Experiments

Experimental flow and data preparation.

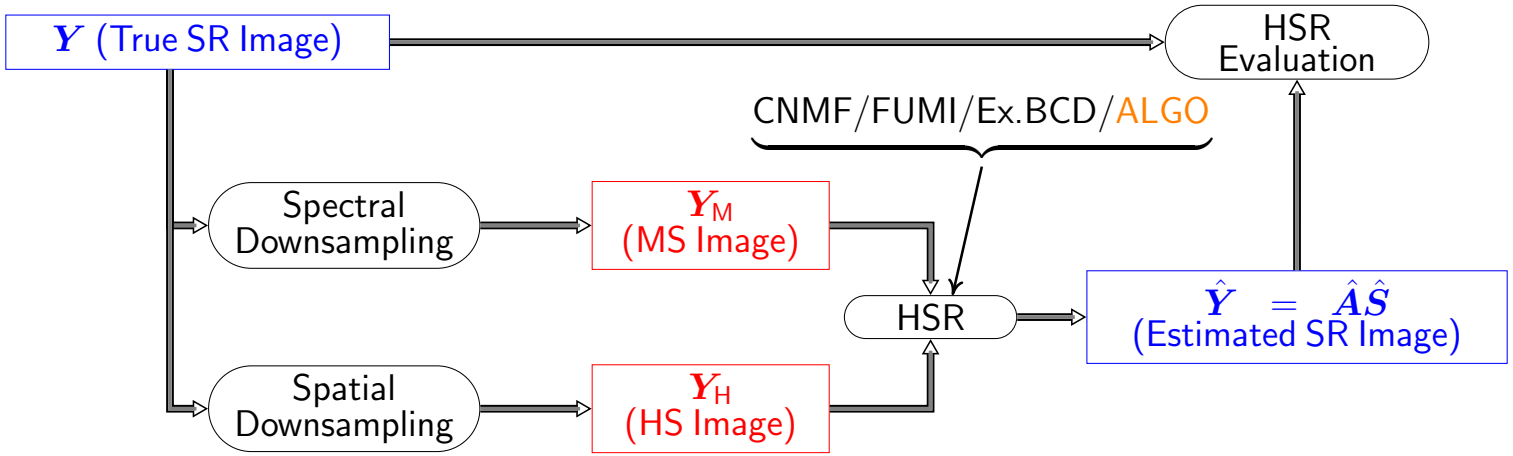


Figure: Wald's Protocol [24] on HS and MS images synthesis and the evaluation of a testing HSR method.

Experiments

Experimental setup:

- use real hyperspectral images as the super-resolution (SR) image;
  - over 100 spectral bands from $0.4\,\mu$m to $3\,\mu$m wavelength.
- generate HS image from the SR image by
  - passing through a Gaussian filter, 2-D mask: 11×11, $\sigma = 1.7$;
  - subsampling on every 4 pixels in vertical and horizontal directions.

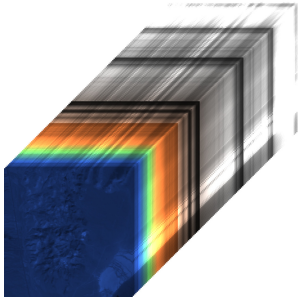
SR image Blurred SR image HS image

Figure: From SR image to HS image, visualized by RGB component.

Experiments

Experimental setup:

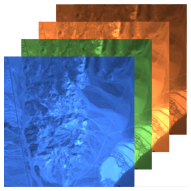
- generate MS image from real SR image by
  - simulate each band of MS band image by taking the average of all band image within the respective passed-band.



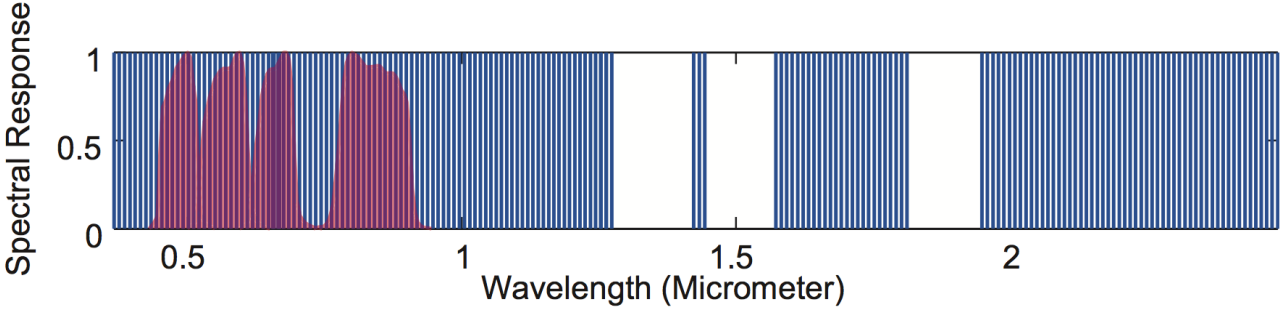
SR image

Wavelength (μm)	Spectral Band Placement
$0.45 - 0.52$	Visible (VIS-B)
$0.52 - 0.60$	Visible (VIS-G)
$0.63 - 0.69$	Visible (VIS-R)
$0.76 - 0.90$	Near Infrared (NIR)

(a) Pass bands of TM sensor in LANDSA 7 Project



MS image



(b) Center wavelengths of HS bands (blue bars) and spectral response of TM sensor (red).

Experiments

Part I: Exact BCD vs ALGO on Cuprite data set.

Implementing Algorithm 1 by adjusting $J = \begin{cases} 100, & \text{Exact BCD;} \\ 1, & \text{ALGO.} \end{cases}$

SNR (dB)	Method	RMSE(dB)	SAM(deg.)	PSNR(dB)	Outer Iter.
40	ALGO-GP	-23.18 ± 0.08	0.88 ± 0.01	44.27 ± 0.13	984.38 ± 50.37
	Ex.BCD-GP	-21.78 ± 0.16	1.19 ± 0.03	42.06 ± 0.22	243.32 ± 14.47
20	ALGO-GP	-16.01 ± 0.04	4.27 ± 0.05	28.77 ± 0.09	195.02 ± 10.27
	Ex.BCD-GP	-14.27 ± 0.13	6.87 ± 0.25	25.28 ± 0.26	54.99 ± 4.19

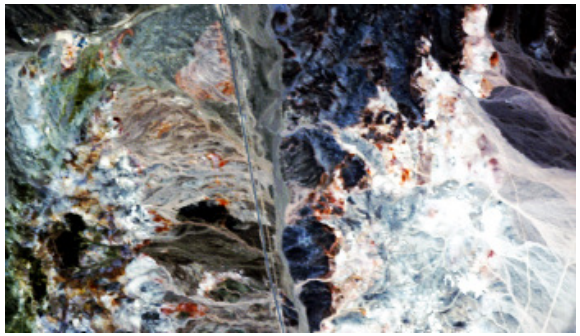
Table: HSR performance of ALGO-GP and Ex.BCD-GP on Cuprite dataset.

SNR (dB)	Method	RMSE(dB)	SAM(deg.)	PSNR(dB)	Outer Iter.
40	ALGO-PG	-23.38 ± 0.02	0.84 ± 0	44.77 ± 0.04	2516.63 ± 98.48
	Ex.BCD-PG	-23.06 ± 0.06	0.91 ± 0.01	44.08 ± 0.1	269.25 ± 20.68
20	ALGO-PG	-16.51 ± 0.02	3.66 ± 0.02	29.88 ± 0.05	475.63 ± 25.43
	Ex.BCD-PG	-15.51 ± 0.04	4.95 ± 0.05	27.72 ± 0.09	80.46 ± 4.28

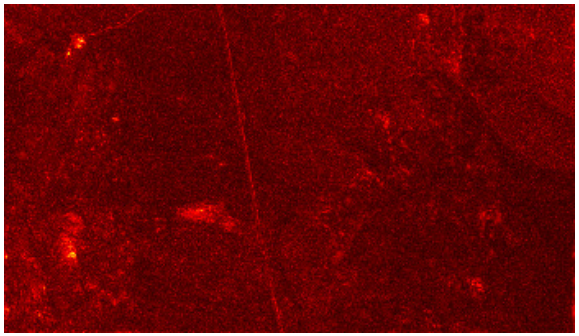
Table: HSR performance of ALGO-PG and Ex.BCD-PG on Cuprite dataset.

Experiments

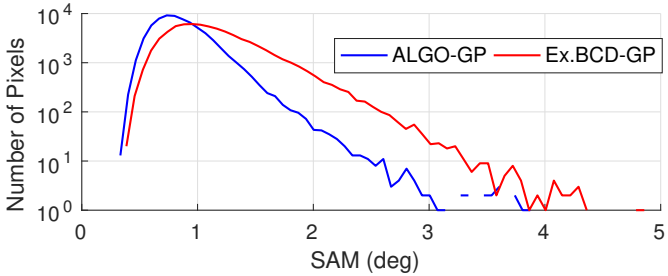
Part 1: Exact BCD vs ALGO on Cuprite data set.



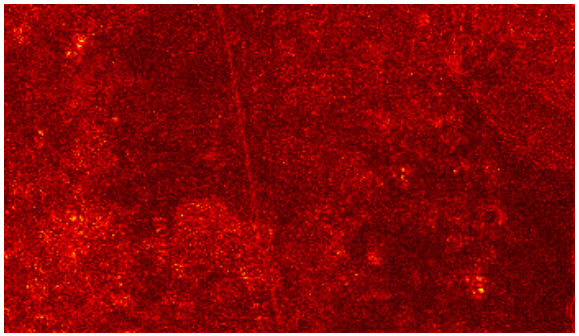
(a) RGB image



(b) SAM map of ALGO-GP



(d) SAM historgram. Blue: ALGO-GP; Red: Ex.BCD-GP.



(c) SAM map of Ex.BCD-GP

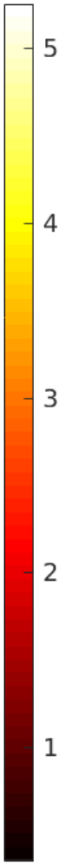


Figure: HSR performance of ALGO-GP and Ex.BCD-GP on Cuprite data set. SNR $= 40$dB; $N = 16$; 100 trials.

Experiments

Part 1: Exact BCD vs ALGO on Synthetic data.

SNR (dB)	Method	RMSE(dB)	SAM(deg.)	PSNR(dB)	Time(sec.)	Outer Iter.
40	ALGO-GP	-15.82 ± 0.38	4.22 ± 0.4	28.9 ± 0.56	13.5 ± 5.85	878.83 ± 133.88
	Ex.BCD-GP	-15.72 ± 0.4	4.37 ± 0.36	28.19 ± 0.53	246.94 ± 74.52	159.92 ± 24.28
20	ALGO-GP	-14.43 ± 0.36	6.64 ± 0.28	22.66 ± 0.46	3.22 ± 1.55	178.06 ± 17.08
	Ex.BCD-GP	-13.13 ± 0.44	9.84 ± 0.22	20.05 ± 0.6	78.48 ± 18.32	50.04 ± 5.78

Table: HSR performance of ALGO-GP and Ex.BCD-GP on synthetic data. $N = 16$; $L = 100 \times 100$; 1000 trials.

SNR (dB)	Method	RMSE(dB)	SAM(deg.)	PSNR(dB)	Time(sec)	Outer Iter.
40	ALGO-PG	-16.86 ± 0.35	3.58 ± 0.37	30.57 ± 0.83	36.28 ± 14.9	3150.36 ± 203.2
	Ex.BCD-PG	-16.39 ± 0.38	3.76 ± 0.38	29.89 ± 0.73	219.61 ± 58.6	206.66 ± 18.57
20	ALGO-PG	-14.86 ± 0.33	5.97 ± 0.32	23.48 ± 0.54	9.99 ± 4.4	758.43 ± 26.36
	Ex.BCD-PG	-13.91 ± 0.4	7.94 ± 0.21	21.66 ± 0.5	97.46 ± 19.34	90.13 ± 3.81

Table: HSR performance of ALGO-PG and Ex.BCD-PG on synthetic data. $N = 16$; $L = 100 \times 100$; 1000 trials.

Experiments

Part 1: Exact BCD vs ALGO
Conclusion: ALGO is much better than Exact BCD in terms of **runtime** while HSR quality is still comparable.

Experiments

Part 2: ALGO using different 1st order gradient methods.

SNR (dB)	Method	SAM(deg.)	PSNR(dB)	Time(sec.)
40	ALGO-GP	4.22 ± 0.4	28.9 ± 0.56	13.5 ± 5.85
	ALGO-BBGP	4.1 ± 0.36	29.13 ± 0.58	17.11 ± 7.58
	ALGO-PG	3.58 ± 0.37	30.57 ± 0.83	36.28 ± 14.9
	ALGO-FISTA	3.07 ± 0.54	32.22 ± 1.55	6.46 ± 2.86
	ALGO-FW	4.7 ± 0.37	27.95 ± 0.51	10.25 ± 4.93
30	ALGO-GP	4.58 ± 0.38	26.79 ± 0.46	6.87 ± 3.2
	ALGO-BBGP	4.51 ± 0.34	26.89 ± 0.48	8.72 ± 4.06
	ALGO-PG	4.18 ± 0.31	27.81 ± 0.64	15.95 ± 7.2
	ALGO-FISTA	3.93 ± 0.48	28.49 ± 1.17	4.41 ± 2.14
	ALGO-FW	5.02 ± 0.36	26.05 ± 0.44	4.43 ± 2.2

Table: HSR performance of ALGO with different $1st$ order methods on synthetic data. $N = 16$; $L = 100 \times 100$; 1000 trials.

Experiments

Part 1: Exact BCD vs ALGO
Conclusion: ALGO is much better than Exact BCD in terms of **runtime** while HSR quality is still comparable.

Part 2: ALGO using different 1st order gradient methods.
Conclusion: ALGO-FISTA has better HSR quality and fewer runtime.

Experiments

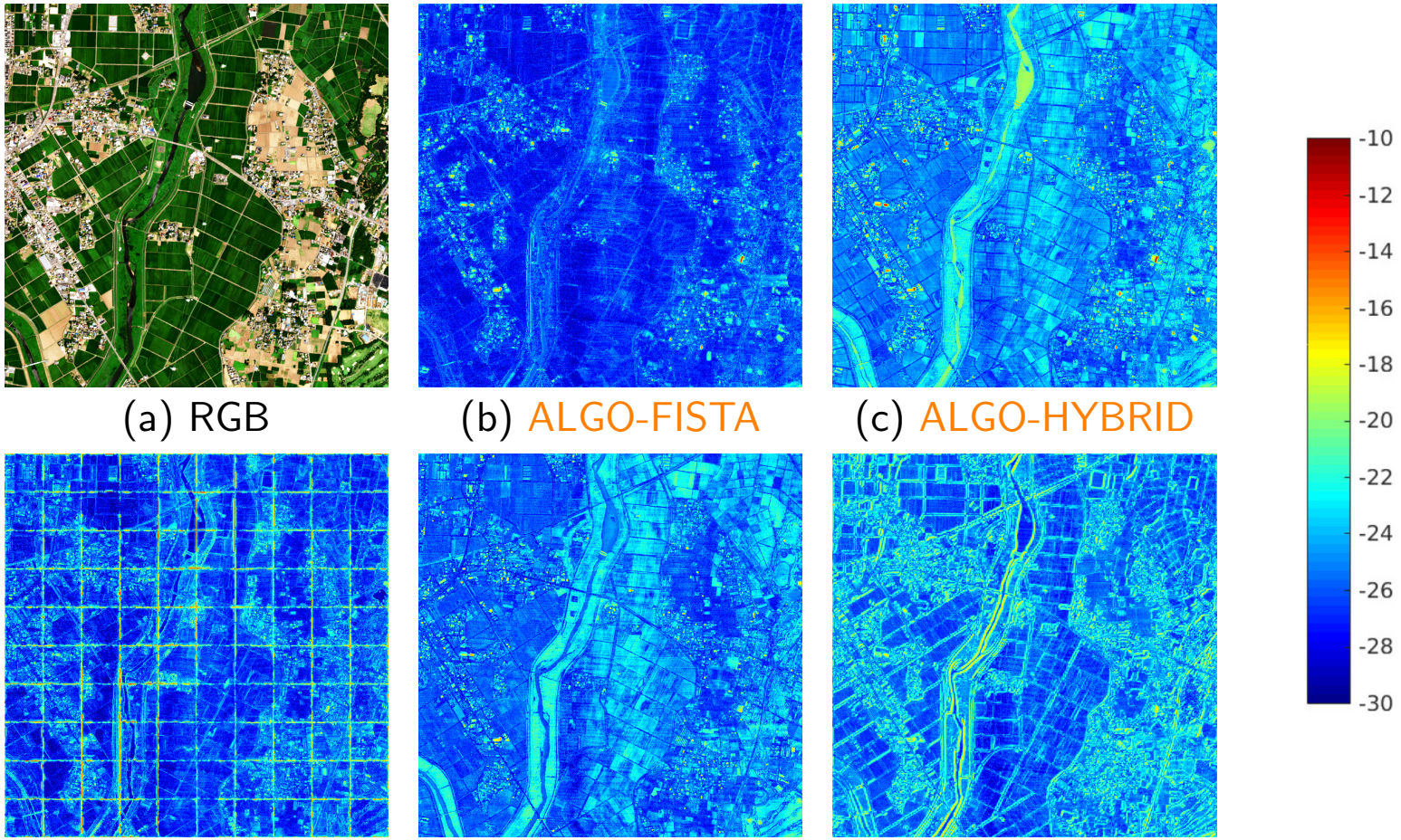
Part 3: ALGO-FISTA vs state-of-the-art algorithms on Chikusei data set (a big image with 1000×1000 pixels).

Concern:

- Chikusei image is too big ($L = 1000^2$!!);
- projection of $\boldsymbol{S}$ onto $\mathcal{S} = \{\boldsymbol{S} \mid \boldsymbol{S} \geq \mathbf{0},\ \mathbf{1}^\top \boldsymbol{S} = \mathbf{1}^\top\}$ is expensive;
- we try to avoid $\Pi_{\mathcal{S}}(\cdot)$ by using FW in the $\boldsymbol{S}$ subproblem and preserving FISTA in the $\boldsymbol{A}$ subproblem. Named "ALGO-HYBRID".

Experiments

Part 3: ALGO vs state-of-the-art algorithms on Chikusei data set (a big image with 1000×1000 pixels).



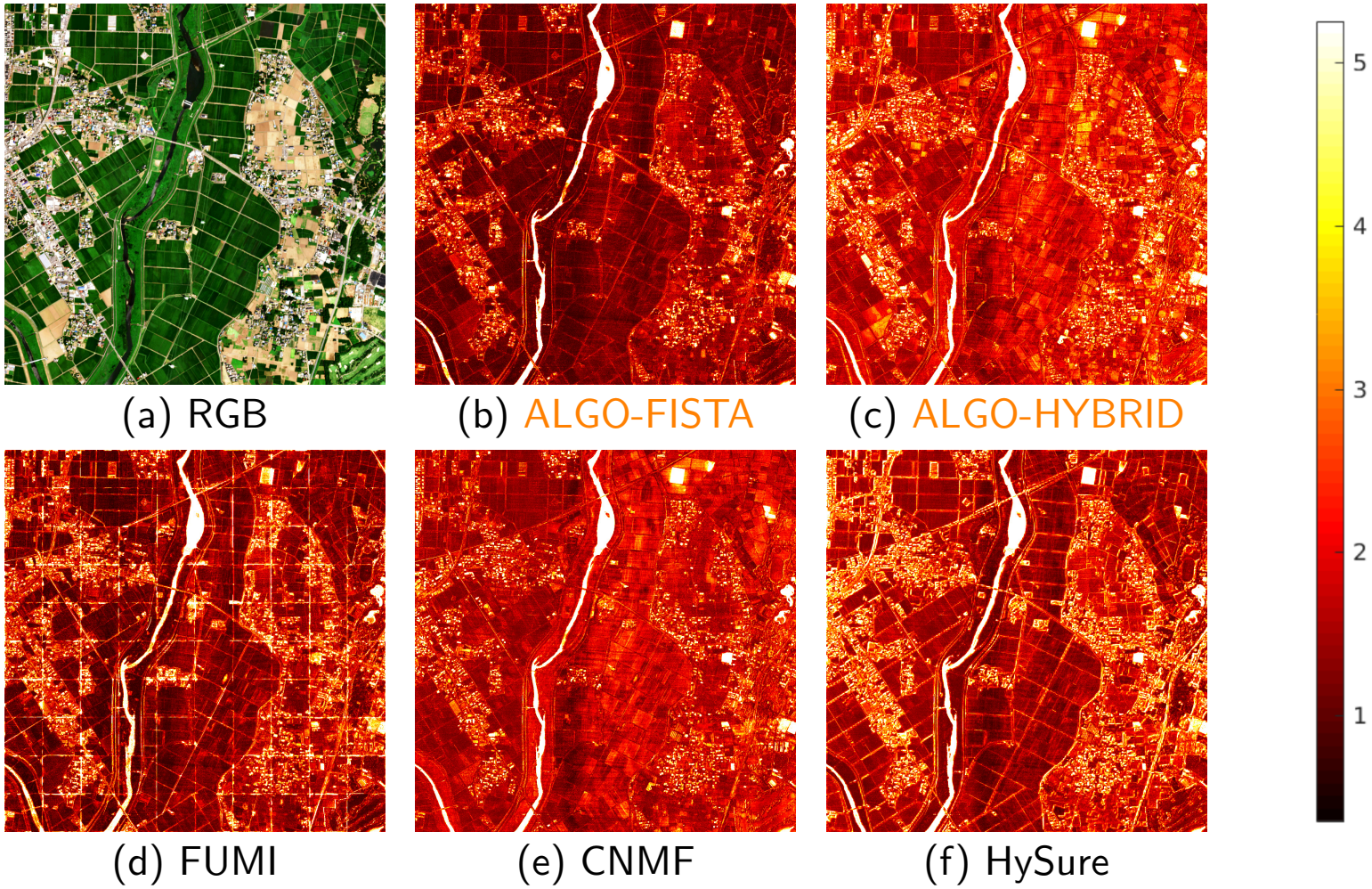
(a) RGB (b) ALGO-FISTA (c) ALGO-HYBRID

(d) FUMI (e) CNMF (f) HySure

Figure: RMSE map of Chikusei data set.

Experiments

Part 3: ALGO vs state-of-the-art algorithms on Chikusei data set (a big image with 1000×1000 pixels).



(a) RGB (b) ALGO-FISTA (c) ALGO-HYBRID

(d) FUMI (e) CNMF (f) HySure

Figure: SAM map of Chikusei data set.

Experiments

Part 3: ALGO vs state-of-the-art algorithms on Chikusei data set (a big image with 1000×1000 pixels).

SNR (dB)	Method	RMSE(dB)	SAM(deg.)	PSNR(dB)	Time(sec.)
40	FUMI	-22.58 ± 0.03	2.07 ± 0.01	39.61 ± 0.08	2744.03 ± 440.45
	CNMF	-23.77 ± 0.02	2.24 ± 0	44.36 ± 0.01	21.19 ± 1.81
	HySure	-23.35 ± 0.11	2.61 ± 0.06	40.47 ± 0.25	39.3 ± 6.36
	ALGO-FISTA	-25.12 ± 0	1.72 ± 0	44.5 ± 0.03	87.23 ± 21.39
	ALGO-HYBRID	-23.08 ± 0.05	2.75 ± 0.04	39.58 ± 0.11	32.5 ± 8.64

Table: HSR performance of ALGO-FISTA and ALGO-HYBRID and FUMI. $N = 9$; $L = 1000 \times 1000$; 100 trials.

Experiments

Part 3: ALGO vs state-of-the-art algorithms on Chikusei data set.
Conclusion:

- ALGO-FISTA has better HSR performance.
- big image processing $\rightarrow$ ALGO-HYBRID can be faster than ALGO-FISTA.
- CNMF is nevertheless the fastest HSR algorithm.

Outline

Introduction and Motivation
- Spectral Signature and Spectral Image
- Applications
- The Instruments
- What is the Problem ?

Hyperspectral Super-resolution (HSR)
- The HSR Problem
- State of the Arts

Fast Algorithm Framework of HSR (proposed)
- Preliminaries
- A Very Fast Algorithm Framework

Experiments

Summary

Appendices and References

Summary

Summary:

- a fast algorithm framework in hyperspectral super-resolution;
- ALGO: an alternating optimization using Inexact BCD with 1st order algorithms;
- ALGO requires fewer runtime compared with FUMI and HySure;
- illustrate ALGO's capability in handling big data sets.

Further discussion:

- the algorithm framework can be extent to similar HSR problem with priori in sparse / smooth abundance $\boldsymbol{S}$;
- direct extension to hyperspectral unmixing (HU)

$$\min_{\boldsymbol{A}\in\mathcal{A},\boldsymbol{S}\in\mathcal{S}} \|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{S}\|_{\mathsf{F}}^2;$$

- great potential in problems of MF form.

Outline

Introduction and Motivation
- Spectral Signature and Spectral Image
- Applications
- The Instruments
- What is the Problem ?

Hyperspectral Super-resolution (HSR)
- The HSR Problem
- State of the Arts

Fast Algorithm Framework of HSR (proposed)
- Preliminaries
- A Very Fast Algorithm Framework

Experiments

Summary

Appendices and References

Appendix: Lee-Seung Multiplicative Update (MU)

return

$$\begin{aligned} \min_{\boldsymbol{W},\boldsymbol{H}} \quad & \|\boldsymbol{V} - \boldsymbol{W}\boldsymbol{H}\|_{\mathsf{F}}^2 \\ \text{s.t.} \quad & \boldsymbol{W} \in \mathbb{R}_+^{m\times n}, \quad \boldsymbol{H} \in \mathbb{R}_+^{n\times \ell}. \end{aligned}$$

Lee-Seung MU

procedure
Input: $\boldsymbol{W}_0 \geq \boldsymbol{0}$, $\boldsymbol{H}_0 \geq \boldsymbol{0}$; Output: $\boldsymbol{W}$, $\boldsymbol{H}$.
    $\boldsymbol{W} \leftarrow \boldsymbol{W}_0$, $\boldsymbol{H} \leftarrow \boldsymbol{H}_0$.
    for $i = 0, 1, 2, \cdots$ **do**
        for $j = 1, 2, \cdots$ **do**
            $\boldsymbol{H} \leftarrow \boldsymbol{H} \cdot \frac{\boldsymbol{W}^\top \boldsymbol{V}}{\boldsymbol{W}^\top \boldsymbol{W}\boldsymbol{H}}$.
        end for
        for $j = 1, 2, \cdots$ **do**
            $\boldsymbol{W} \leftarrow \boldsymbol{W} \cdot \frac{\boldsymbol{V}\boldsymbol{H}^\top}{\boldsymbol{W}\boldsymbol{H}\boldsymbol{H}^\top}$.
        end for
        $\boldsymbol{W}$, $\boldsymbol{H}$.
    end for
end procedure

Appendix: HSR as QP

return

Let $\boldsymbol{y}_1$, $\boldsymbol{y}_2$ and $\boldsymbol{x}$ be the vectorization of $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$ and $\boldsymbol{X}$.

$$
\begin{aligned}
& \min_{\boldsymbol{X}} \|\boldsymbol{Y}_1 - \boldsymbol{A}_1 \boldsymbol{X} \boldsymbol{B}_1\|_{\mathsf{F}}^2 + \|\boldsymbol{Y}_2 - \boldsymbol{A}_2 \boldsymbol{X} \boldsymbol{B}_2\|_{\mathsf{F}}^2 \\
= \quad & \min_{\boldsymbol{x}} \|\boldsymbol{y}_1 - \left[\boldsymbol{B}_1^\top \otimes \boldsymbol{A}_1\right] \boldsymbol{x}\|_2^2 + \|\boldsymbol{y}_2 - \left[\boldsymbol{B}_2^\top \otimes \boldsymbol{A}_2\right] \boldsymbol{x}\|_2^2 \\
= \quad & \min_{\boldsymbol{x}} \left\| \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{B}_1^\top \otimes \boldsymbol{A}_1 \\ \boldsymbol{B}_2^\top \otimes \boldsymbol{A}_2 \end{bmatrix} \boldsymbol{x} \right\|_2^2 \\
= \quad & \min_{\boldsymbol{x}} \quad \boldsymbol{x}^\top \boldsymbol{P} \boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + r
\end{aligned}
$$

References I

[1] R. F. Kokaly, R. N. Clark, G. A. Swayze, K. E. Livo, T. M Hoefen, N. C.; Pearson, R. A. Wise, W. M. Benzel, H. A. Lowers, R. L. Driscoll, and A. J. Klein, "USGS Spectral Library Version 7: U.S. Geological Survey Data Series 1035," `http://speclab.cr.usgs.gov/`, 2017.

[2] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral Unmixing Overview: Geometrical, Statistical, and Sparse Regression-Based Approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

[3] F. A. Kruse, "Comparison of AVIRIS and Hyperion for Hyperspectral Mineral Mapping," in *Proceedings of the 11th JPL Airborne Geoscience Workshop, 4-8 March 2002, Pasadena, California*, 2002.

[4] "AVIRIS Data - Ordering Free AVIRIS Standard Data Products.," [Online]. Available"`http://aviris.jpl.nasa.gov/html/aviris.freedata.html`, 2016, [Accessed: 2016-05-01].

References II

[5] P. E. Dennison, , D. A. Roberts, and L. Kammer, "Wildfire Detection for Retrieving Fire Temperature from Hyperspectral Data," in *American Society for Photogrammetry and Remote Sensing - ASPRS Annual Conference 2008 - Bridging the Horizons: New Frontiers in Geospatial Collaboration*, 2008, vol. 1, pp. 139–146.

[6] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging Spectrometry for Earth Remote Sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, 1985.

[7] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and Multispectral Data Fusion: A Comparative Review of the Recent Literature," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, June 2017.

[8] "Landsat 7 Science Data Users Handbook," `http://landsat.gsfc.nasa.gov/wp-content/uploads/2016/08/Landsat7_Handbook.pdf`, 2016, Accessed: 2016-09-29.

[9] J. A. Barsi, K. Lee, G. Kvaran, B. L. Markham, and J. A. Pedelty, "The Spectral Response of the Landsat-8 Operational Land Imager," *Remote Sensing*, vol. 6, no. 10, pp. 10232–10251, 2014.

References III

[10] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J. Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral Pansharpening: A Review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 27–46, Sept 2015.

[11] W. K. Ma, J. M. Bioucas-Dias, T. H. Chan, N. Gillis, P. Gader, A. J. Plaza, A. Ambikapathi, and C. Y. Chi, "A Signal Processing Perspective on Hyperspectral Unmixing: Insights from Remote Sensing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 67–81, Jan 2014.

[12] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528–537, Feb 2012.

[13] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J. Y. Tourneret, M. Chen, and S. Godsill, "Multiband Image Fusion Based on Spectral Unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7236–7249, Dec 2016.

[14] D. D. Lee and H. S. Seung, "Learning the Parts of Objects by Non-negative Matrix Factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct 1999.

References IV

[15] E. Wycoff, T. H. Chan, K. Jia, W. K. Ma, and Y. Ma, "A Non-Negative Sparse Promoting Algorithm for High Resolution Hyperspectral Imaging," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 1409–1413.

[16] M. Simes, J. BioucasDias, L. B. Almeida, and J. Chanussot, "A Convex Formulation for Hyperspectral Image Superresolution via Subspace-Based Regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3373–3388, June 2015.

[17] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific Belmont, 1999.

[18] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141, 1988.

[19] N. Parikh and S. Boyd, "Proximal Algorithms," *Foundations and Trends in Optimzation*, vol. 1, no. 3, pp. 127–239, Jan 2014.

[20] Y. Nesterov, *Introductory Lectures on Convex Optimization : A Basic Course*, Applied optimization. Kluwer Academic Publ., Boston, Dordrecht, London, 2004.

[21] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

References V

[22] M. Frank and P. Wolfe, "An Algorithm for Quadratic Programming," *Naval Research Logistics Quarterly*, vol. 3, pp. 95–110, March-June 1956.

[23] N. Yokoya and A. Iwasaki, "Airborne Hyperspectral Data over Chikusei," Tech. Rep. SAL-2016-05-27, Space Application Laboratory, University of Tokyo, Japan, May 2016.

[24] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of Satellite Images of Different Spatial Resolutions: Assessing the Quality of Resulting Images," *hotogrammetric Engineering and Remote Sensing*, vol. 63, no. 6, pp. 691–699, 1997.